OXFORD
UNIVERSITY PRESS

Archives of Clinical Neuropsychology 38 (2023) 699–713

Archives
of
CLINICAL
NEUROPSYCHOLOGY

Remote, Computerised Cognitive Assessment for Breast Cancer- and Treatment-Related Cognitive Dysfunction: Psychometric Characteristics of the Cogsuite Neurocognitive Battery

James C. Root[1,*], Alexandra M. Gaynor[1,2,3], Anam Ahsan[1], Duane Jung[4], Elizabeth Schofield[1], Elizabeth Ryan[1], Yuelin Li[1], Tim A. Ahles[1]

[1]*Department of Psychiatry and Behavioral Sciences, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA*
[2]*Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University, New York, NY 10027, USA*
[3]*Cognitive Neuroscience Division, Department of Neurology, Columbia University, New York, NY 10027, USA*
[4]*Enformia Inc., Davidson, NC 28031, USA*

*Corresponding author at: Neurocognitive Research Laboratory, Department of Psychiatry and Behavioral Sciences, Memorial Sloan Kettering Cancer Center, 641 Lexington Ave., 7th Floor, New York, NY 10022, USA.
E-mail address: rootj@mskcc.org (J.C. Root).

Accepted 19 December 2022

Abstract

Objective: Cancer-related cognitive dysfunction (CRCD) is a significant concern for breast cancer survivors. The Cogsuite battery was developed to improve sensitivity to CRCD with the use of cognitive experimental measures, clarify specific cognitive processes impacted and to be capable of being administered either in-office or remotely.
Methods: In sum, 357 breast cancer survivors and non-cancer controls completed the Cogsuite Battery in-office ($n = 76$) or remotely ($n = 281$). Measure validity, sensitivity to demographic factors, correlations with standard neuropsychological measures and intercorrelations of Cogsuite variables were assessed. Test-retest reliability was evaluated in-office ($n = 24$) and remotely ($n = 80$).
Results: Test-retest reliability for most variables assessed was adequate to strong. Internal validity, as indicated by the confirmation of expected condition effects within each measure, was established for all measures. Assessment of external validity found age, but not education, was a significant predictor in the majority of measures. Assessment of criterion validity found that Cogsuite variables were correlated with standard measures in psychomotor speed, working memory and executive function, but not associated with self-reported cognition or mood.
Conclusions: Cogsuite is reliable and valid, and is sensitive to the effects of increasing age on cognition. The addition of the Cogsuite battery to standard assessment may improve sensitivity to CRCD and identify underlying processes that may be affected. Remote use of the Cogsuite battery in appropriate settings will lessen the burden for providers, researchers and survivors in research and clinical contexts.

Keywords: Breast cancer; Cognitive impairment; Survivorship; Cancer-related cognitive dysfunction

Introduction

Given US population demographics and improvement in cancer survival, the US will see an unprecedented increase in individuals living with a history of cancer. Cancer-related cognitive dysfunction (CRCD) in non-central nervous system cancers is a significant concern for cancer survivors, particularly for older cancer survivors. A substantial body of research in breast cancer survivors finds declines in cognition by self-report, as measured by standard assessments around the time of treatment,

https://doi.org/10.1093/arclin/acac111 Advance Access publication 18 January 2023

in short intervals after, as well as in long-term survivors (Ahles & Root, 2018). Standard neuropsychological measures have revealed alterations in attention, learning, psychomotor speed and executive function from months to years post-treatment (Ahles & Root, 2018; Deprez et al., 2014; Dijkshoorn et al., 2021; Dyk, Crespi, Petersen, & Ganz, 2020; Janelsins et al., 2018; Mandelblatt et al., 2018; Root, Andreotti, Tsu, Ellmore, & Ahles, 2016; Wefel, Lenzi, Theriault, Davis, & Meyers, 2004). In the case of breast cancer, several potential exposures and mechanisms have been considered, including biological stress related to surgery, inflammation associated with cancer and treatment, cytotoxic effects of chemotherapy and changes in hormonal milieu associated with hormonal therapies for control of recurrence in hormone-receptor positive tumors. The majority of breast cancer survivors expect that they will return to previous responsibilities, but report difficulties returning to household and family management, the need for more time and effort in their work, increased frustration, financial impact, as well as decreased quality of life related to CRCD (Boykoff, Moieni, & Subramanian, 2009; Munir, Burrows, Yarker, Kalawsky, & Bains, 2010; Myers, 2010; Selamat, Loh, Mackenzie, & Vardy, 2014; Von Ah, Habermann, Carpenter, & Schneider, 2013). Recent work has found associations of work ability with both self-reported and objective cognitive dysfunction (Klaver et al., 2022).

Given the sometimes subtle presentation of CRCD in breast cancer and other non-central nervous system cancers, there has been growing interest in developing cognitive-experimental measures to be used in addition to standard measures to improve sensitivity to cognitive dysfunction and identify specific processes that might be implicated (Horowitz, Suls, & Trevino, 2018; Trevino et al., 2021). Similar issues have arisen in other clinical areas leading to the development of cognitive-experimental measures to better assess cognition associated with clinical syndromes (NIH Examiner; Kramer et al., 2014) and schizophrenia (CNTRICS; Barnett et al., 2010). The NCI requested proposals to develop online, remotely deliverable cognitive measures informed by cognitive-experimental methods and neuroscience theory for assessment of CRCD (NIH/NCI 343: An Electronic Platform for Cognitive Assessment in Cancer Patients). Through that mechanism, we have developed a battery of measures for use in assessing CRCD, followed by data collection to assess psychometric characteristics of the included measures. The Cogsuite Battery (https://www.enformia.com/cogsuite/) (Supplementary Fig. S1) is a platform of web-based neurocognitive measures designed for unsupervised, in-office or remote assessment of cognition in cancer patients and survivors. Although developed for the assessment of CRCD, the battery is also anticipated to be of use in other cognitive syndromes that can be similarly subtle in nature (e.g., remitted multiple sclerosis, attention deficit disorder, post-concussive syndrome). The battery includes seven measures that can be administered together or modularly, assessing attentional function, inhibitory control, working memory, visuospatial function, phonemic and semantic fluency, psychomotor speed and motor speed; web-based, self-report measures are also included for assessment of mood and cognitive function and can be edited or expanded to include other self-report inventories as needed. An administrator portal is included for assigning, delivering and tracking the completion of cognitive tasks, data summarisation and visualisation, and review of individual- or group-level performance characteristics (Supplementary Fig. S2).

In choosing and developing the cognitive tasks included, we reviewed the literature for measures that demonstrated sensitivity to CRCD and could be implemented and administered online and remotely, as discussed below.

Attention—Attention Network Test (ANT)

We were interested in measuring specific aspects of attention and working memory due to previous research that suggests the presence of initial learning deficits in learning and memory measures in survivors, with compensation through repeated learning trials and normative retention and recall (Gaynor et al., 2021; Root et al., 2015; Root, Andreotti, Tsu, Ellmore, & Ahles, 2016). The ANT paradigm (Fan et al., 2009) was chosen for its ability to specifically measure three components of attention: Alerting, Orienting and Executive Control. The central task is to identify the direction of the center stimulus (the target) in an array of either congruent or incongruent stimuli. Briefly, four cue conditions preceding the flanker array offer varying information regarding when and where the flanker will be presented: The no-cue condition offers no temporal or spatial information; the double cue condition offers temporal information; the valid cue offers both temporal and spatial information; and the invalid cue offers correct temporal information and incorrect spatial information. The differences in cue and target conditions allow for identifying alterations in specific attentional networks. Previous work using the attention network test found specific deficits in the alerting and executive control networks in survivors, with intact performance in orienting (Chen et al., 2014). Including this measure may resolve specific processes that are affected in CRCD.

Working Memory—N-back Test

Attention and working memory have been widely studied in post-treatment cancer survivors using the N-back test (Conroy et al., 2013; de Ruiter et al., 2011; Ferguson, McDonald, Saykin, & Ahles, 2007; Kesler, Bennett, Mahaffey, & Spiegel, 2009; McDonald, Conroy, Ahles, West, & Saykin, 2012). Briefly, the task consists of presentation of single letters during which the individual determines if the current letter is a match according to the following criteria: 0-back—the letter matches a specified letter; 1-back—the letter is the same as the letter immediately preceding it; 2-back—the letter is the same as the one two letters before; 3-back—the letter is the same as the one three letters before. Each condition imposes an increasing working memory load. The N-back task (Gevins & Cutillo, 1993) has demonstrated sensitivity to differences in prefrontal brain activation in survivors (McDonald, Conroy, Ahles, West, & Saykin, 2010; McDonald, Conroy, Ahles, West, & Saykin, 2012; Root et al., 2021; Scherling, Collins, Mackenzie, Bielajew, & Smith, 2011), and neurocognitive studies have demonstrated sensitivity of the N-back to the effects of childhood cancer treatment and age at treatment in adult survivors (Luxton et al., 2014). Significantly, functional changes have been associated with underlying structural changes following treatment, suggesting that N-back performance may be sensitive to structural changes in volume and/or density in survivors (McDonald, Conroy, Ahles, West, & Saykin, 2010, 2012; McDonald, Conroy, Smith, West, & Saykin, 2013).

Executive Functioning—Stop Signal Task

The stop-signal task (Logan, Cowan, & Davis, 1984) is a measure of sustained attention, cognitive control and response inhibition, and has been demonstrated to be sensitive to a subset of psychiatric and cognitive disorders, with particular sensitivity in attention deficit disorder (Lipszyc & Schachar, 2010). Briefly, the task presents a series of stimuli in a choice reaction time format and on a subset of trials the target is followed by a stop-signal indicating that a response should be withheld, with the delay of the stop signal determined by previous success or failure on the last stop-signal trial. Tasks of response inhibition have been previously used in studies of CRCD and have shown pre-treatment differences in fMRI activation (Scherling, Collins, Mackenzie, Bielajew, & Smith, 2012) and post-treatment differences in event-related potentials and regional functional recruitment (Chao et al., 2012; Simo, Gurtubay-Antolin, Vaquero, Bruna, & Rodriguez-Fornells, 2018), suggesting sensitivity to subtle changes in executive function related to cancer and treatment. Together with the frontal-executive measures that include the N-back working memory task and the modified attention network task described above, the stop-signal task was included to further clarify any potential weakness in frontal-executive systems specific to response inhibition.

Visuospatial Ability—Mental Rotation

The mental rotation task (Shepard & Metzler, 1971) presents two visual stimuli at varying rotations from each other and requires mentally rotating one of the figures to determine if the two shapes are the same or different. Changes in visuospatial abilities may be particularly important in the assessment of individuals on endocrine therapies, and specifically in androgen deprivation therapy used for the treatment of men with prostate cancer. Visuospatial abilities have been found to be associated with normally varying testosterone levels in men both with regard to matching accuracy and reaction time (Hooven, Chabris, Ellison, & Kosslyn, 2004). In men with prostate cancer treated with androgen deprivation therapy, mental rotation errors were increased compared with men not treated (Cherrier, Aubin, & Higano, 2009).

Verbal Abilities—Verbal Fluency Task

The verbal fluency task included in this battery requires typed generation of words beginning with specific letters (F, A, S) or belonging in a given semantic category (Animals) (Lezak, 1995). Survivors' report of changes in verbal abilities, particularly in word finding and word substitution is common. Findings from a recent meta-analysis in cancer survivors confirm that verbal abilities, including vocabulary, word finding and word generativity, are particularly affected following treatment (Jim et al., 2012). Results from structural and functional neuroimaging studies support a word generativity weakness, as suggested by a pattern of change in structure and function in anterior, dorsolateral prefrontal cortex that is associated with support of efficient verbal expression (McDonald, Conroy, Ahles, West, & Saykin, 2010, 2012; McDonald, Conroy, Smith, West, & Saykin, 2013). In a longitudinal analysis from our group (Ahles et al., 2010), chemotherapy-exposed patients exhibited an altered trajectory in verbal fluency performance from pre- to post-treatment compared with unexposed patients and healthy controls.

Psychomotor Speed—Coding

The coding task requires determining whether successive pairings of two stimuli match a pre-determined key. Although we have included in the online battery measures of simple and choice reaction time, higher order processing speed, in which the speed of processing is determined by more complex decision processes, is assessed with this task. Psychomotor slowing has been demonstrated to be a sensitive but non-specific marker of cognitive dysfunction associated with several neurological disorders and with normal aging (Rosano et al., 2016; Rosano, Newman, Katz, Hirsch, & Kuller, 2008). Psychomotor speed has been found to be affected in cancer survivors, with performance on speeded measures correlated with white matter integrity as assessed by DTI, and significant slowing at both pre- and post-treatment timepoints (Deprez et al., 2011; Lyon et al., 2016).

Motor Function—Finger Tapping

The task requires speeded repetitive finger tapping over brief intervals with each hand. Simple motor dexterity and speed has been found to be a sensitive and non-specific indicator of lateralised hemispheric involvement (Robinson, Fitts, & Kraft, 1990), white matter integrity (Zhai et al., 2020) and cognitive dysfunction (Suzumura et al., 2021). Previous studies have found evidence of decreased motor functioning in chemotherapy-treated survivors. In addition to primary CNS dysfunction, this measure will also be sensitive to peripheral dysfunction associated with certain chemotherapy-induced neuropathies.

Self-report Measures

PROMIS Anxiety: This is an eight-item self-report questionnaire rating symptoms of anxiety on a scale of 1 (never) to 5 (always) over the past 7 days. **PROMIS Depression:** This is an eight-item self-report questionnaire rating symptoms of depression on a scale of 1 (never) to 5 (always) over the past seven days. **PROMIS Cognitive Function**: This is an eight-item self-report questionnaire rating cognitive function on a scale of 1 (never) to 5 (very often). **PROMIS Cognitive Abilities:** This is an eight-item self-report questionnaire rating cognitive function on a scale of 1 (not at all) to 5 (very much) (https://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis/list-of-adult-measures).

We previously reported results of an initial analysis documenting group differences between a cohort of breast cancer survivors and non-cancer controls (Gaynor et al., 2022b), indicating sensitivity of the battery to CRCD. Here we report data for initial validation and psychometric characterisation of the Cogsuite Battery in 357 participants including non-cancer controls and breast cancer survivors. We hypothesised that we would see expected condition effects within each measure, that measures would demonstrate adequate test-retest reliability and that measures would be sensitive to demographic variables such as age and education. Test-retest reliability, condition effects and intercorrelations between individual measures within the Cogsuite battery and with standard neuropsychological measures and self-report measures are reported. We initially examined group by condition effects and found no interaction effects that would indicate that the measures and conditions worked differently based on group status. Therefore, the sample is combined to examine the psychometric properties of Cogsuite.

Methods

Measure Identification and Selection

In the initial proposal for Phase 1 of the project, we identified measures that met criteria of the RFA, specifically, identification of cognitive experimental paradigms informed by neuroscience theory that would likely be sensitive to CRCD. In addition to demonstrated sensitivity to CRCD in previously published literature, we also sought out paradigms that (1) demonstrated adequate reliability, (2) could potentially shed light on aspects of cognition not captured in standard measures but that may be related to performance on those measures, (3) from clinical judgment and use of similar measures in a clinical setting, would likely be sensitive to CRCD and (4) were able to be delivered remotely, unsupervised, and in a timely and efficient manner. The primary author (JCR) was responsible for deciding the final list of included measures. Measures that meaningfully discriminated cancer from control groups were identified as potentially sensitive and examined for inclusion in this battery; the literature considered is cited in the introduction and in the description of each measure. The final selected measures were included in the original grant submission to NCI, where the proposal was reviewed, and are in the Cogsuite battery currently. Usability testing, conducted at an independent usability lab, was conducted where the interface, instructions and browser interaction were modified in response to the user feedback for both the assessment portal as well as the administrator portal.

Participants

Female breast cancer survivors ($N = 174$) were recruited through the Army of Women as well as the survivorship clinic at Memorial Sloan Kettering Cancer Center (MSKCC). Female non-cancer controls ($N = 183$) were recruited through the Army of Women and community advertisement. Inclusion and exclusion criteria followed our methods in previous, NIH-funded survivorship studies. Criteria were developed to limit contributions from other comorbid psychiatric or neurological disorders, and to ensure relative homogeneity of non-cancer control and breast cancer groups. For medical and other history, all treatment and other demographic information was derived from self-report of survivors and controls. Survivors were eligible if they were 40–65 years old, had a history of AJCC stages 0–3 breast cancer, and were between 6 months and 10 years post-treatment, either on current hormonal therapy or not. Survivors were excluded if they had a history of another type of cancer, except non-melanoma/basal cell skin cancer/squamous cell skin carcinoma, or early-stage secondary cancer treated only with surgery. Controls were eligible if they were females aged 40–65 years, with no history of cancer except non-melanoma/basal cell skin cancer/squamous cell skin carcinoma, or early-stage secondary cancer treated only with surgery. For both groups, inclusion criteria also included the following: (1) a score of <11 on the Blessed Orientation-Memory-Concentration Test (BOMC), (2) if currently taking psychoactive medications (excluding gabapentin) on a daily basis, the dose must have been stable for at least two months prior to enrollment; and (3) per-self report, English fluency of "well" or "very well," and having a reasonable comprehension of the study conversation in the opinion of the research staff. For both groups, exclusion criteria included (1) the diagnosis of neurodegenerative disorder that affects cognitive function; (2) history of stroke or head injury resulting in a structural lesion on neuroimaging; (3) persistent cognitive difficulties impacting work or daily life or required cognitive rehabilitation; (4) fine motor/motor impairments that interfere with the participant's ability to use a keyboard; (4) diagnosis of a Schizophrenia Spectrum Disorder, substance use disorder, Bipolar Disorder or Schizotypal personality disorder; (5) self-reported visual or auditory impairment that would preclude ability to complete the assessments; and (6) use of methotrexate or rituximab for rheumatoid arthritis, psoriasis or Crohn's disease, or cyclophosphamide for Lupus. The two cohorts were matched such that they were comparably distributed in age and education. All methods and procedures were approved by the Institutional Review Board at MSKCC.

Procedures

This data collection consisted of (1) participants assessed in-person with both Cogsuite and standard neuropsychological assessment ($n = 76$), and (2) a larger subset assessed remotely with the Cogsuite battery, excluding standard, performance-based neuropsychological measures ($n = 281$). The first half of participants recruited was selected to undergo repeated testing on average 40 days later, yielding a total of 104 paired datasets to assess test-retest reliability; 24 participants were assessed with the Cogsuite battery twice in-office, and 80 participants were assessed twice remotely.

In-office Assessment Measures

For in-office visits, all participants were administered a standard neuropsychological battery to test multiple cognitive domains: premorbid cognition—WRAT Reading Test (Wilkinson & Robertson, 2006); psychomotor speed—Trail-making Part A and B (Army Individual Test Battery, 1944), WAIS-III Digit Symbol Subtest (Kaufman & Lichtenberger, 1999); language—Controlled Oral Word Association Test (Lezak, 1995), Category Fluency Test (Lezak, 1995); attention—Digit Span (Stern & White, 2003), D-KEFS Color-Word Interference Test (Delis, Kaplan, & Kramer, 2001); and memory—Hopkins Verbal Learning Test (Benedict, Schretlen, Groninger, Dobraski, & Shpritz, 1996), Brief Visuospatial Memory Test (Benedict, Schretlen, Groninger, Dobraski, & Shpritz, 1996). Self-report measures were used to assess mood—Center for Epidemiologic Studies Depression (Radloff, 1977), State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970)), fatigue—Fatigue Symptom Inventory (Hann et al., 1998); and perceived cognitive function—Sensory Gating Inventory (Hetrick, Erickson, & Smith, 2012). Brief questionnaires were also used to capture sociodemographic information, general health history and experience with computerised cognitive training. The Cogsuite battery was administered at the same study visit with research staff located in a separate room to avoid distraction and allow for unsupervised administration as intended.

Remote Assessment Measures

For remote assessments, the same Cogsuite battery, self-report measures, and self-reported demographic and health history questionnaires were administered online with participants in their home environments. Standard neuropsychological measures were not administered in remote assessments. For remote administration, self-report data were captured online via REDcap.

Data Analyses

Cogsuite data were transformed prior to analyses using Box-Cox transformation so that they more closely resemble a normal distribution for statistical analyses. All available data are reported, and no participant-level data were omitted for any pre-defined performance characteristics. Raw data were used for descriptive statistics and plots. The performance on each measure was first examined to confirm that expected condition effects and associations of demographic characteristics with reaction time and/or

Table 1. Demographic characteristics

	Control ($N = 183$)	Survivor ($N = 174$)	P-value
Age			
Mean (SD)	55.9 (7.02)	55.1 (6.94)	0.3
Median [Min, Max]	57.0 [39.0, 65.0]	56.0 [40.0, 65.0]	
Education			
High school/GED	8 (4.4%)	11 (6.3%)	0.32
Associate degree/some college	19 (10.4%)	23 (13.2%)	
Vocational/technical school	1 (0.5%)	5 (2.9%)	
Bachelor's degree	72 (39.3%)	61 (35.1%)	
Advanced degree	83 (45.4%)	74 (42.5%)	
Handedness			
Right	164 (89.6%)	159 (91.4%)	0.83
Left	17 (9.3%)	13 (7.5%)	
Ambidextrous	2 (1.1%)	2 (1.1%)	
Marital Status			
Married	130 (71.0%)	122 (70.1%)	0.75
Domestic partnership	5 (2.7%)	10 (5.7%)	
Widowed	3 (1.6%)	2 (1.1%)	
Divorced	18 (9.8%)	19 (10.9%)	
Separated	1 (0.5%)	1 (0.6%)	
Never married	26 (14.2%)	20 (11.5%)	
Employment Status			
>32 hours/week	100 (54.6%)	84 (48.3%)	0.48
<32 hours/week	21 (11.5%)	22 (12.6%)	
Not working	62 (33.9%)	68 (39.1%)	
Racial Identity			
American Indian, Native Alaska	1 (0.5%)	0 (0%)	<0.01
Asian, Native Hawaiian, other	2 (1.1%)	11 (6.3%)	
Black, African-American	4 (2.2%)	16 (9.2%)	
Other	2 (1.1%)	4 (2.3%)	
Refused	1 (0.5%)	2 (1.1%)	
White, Caucasian	173 (94.5%)	141 (81.0%)	
Ethnicity			
Non-Hispanic or Latino	174 (95.1%)	165 (94.8%)	1
Hispanic or Latino	8 (4.4%)	8 (4.6%)	
Missing	1 (0.5%)	1 (0.6%)	
Arm			
In Office	22 (12.0%)	54 (31.0%)	<0.01
Remote	161 (88.0%)	120 (69.0%)	

accuracy were exhibited using repeated measure and univariate ANCOVAs, with age and education as covariates. Pearson's correlation coefficient was used to assess intercorrelations between sets of measures (Cogsuite, standard neuropsychological assessments and self-reported measures) as well as test-retest reliability. All data were prepared and visualised using the statistical language R (Version 1.4.1717) and analyses were conducted using SPSS Version 27.0 (IBM).

Results

Table 1 presents baseline characteristics of the sample ($n = 357$); 21% completed assessments in-office; all others completed assessments remotely online. Participants were largely White (88%), not Hispanic (95%), average age 56 years old (range 39–65), married (71%), with a college education or advanced degree (81%). Half (52%) currently worked outside the home at least 32 hours per week.

Acceptability and Tolerability

Initial useability assessment was conducted in Phase 1 of this project in an independent useability lab where the final adjustments to interface elements and task instructions were made. In the current data collection, acceptability and tolerability were assessed by the number of participants who were assigned, started and completed the battery compared with the number of participants who were assigned and started the battery but did not complete. The Cogsuite battery exhibits strong acceptability and tolerability. All in-office participants completed the entire battery successfully. Of 288 remote participants, 281 (98%) completed the entire battery successfully, whereas 7 did not.

Reliability

The average test-retest interval was 39 days and ranged from 23 to 70 days between sessions (SD = 7.6 days). Test-retest reliability ranged from 0.37 (Stop Signal Go Accuracy) to 0.89 (Finger Tapping Lag Right) (Fig. 1). Of 20 variables assessed, 5 exhibited "fair" test-retest reliability (0.40–0.59), 6 exhibited "moderate" reliability (0.60–0.74) and 8 exhibited "strong" reliability (.75 or higher) (Chan, 2003). Although other variables in the Stop Signal Delay task exhibited fair to adequate reliability, one variable, Stop-signal Go Accuracy, did not exhibit appropriate reliability (<0.40); secondary analysis of this variable revealed severe restriction in range (time 1 kurtosis = 22.8; time 2 kurtosis = 13.5) that potentially affected the observed correlation. Other variables with lower reliability including ANT accuracy and N-back accuracy tended to exhibit similar restrictions in range due to ceiling effects in performance.

Internal and External Validity

Below are summaries of internal and external validity assessments. Internal validity addressed whether each measure performed as expected under the current study design. For instance, condition effects were expected to address validity of the administration given alterations in visual elements of the tasks and remote administration in the majority of cases (results shown in Fig. 2a–h). External validity examined whether the tests performed as expected by baseline characteristics that are known to be associated with cognitive performance, here consisting of age and education. Additionally, criterion validity was sought in the intercorrelations between Cogsuite measures, standard neuropsychological measure and self-reported cognitive performance. Although we use the term criterion validity to describe this analysis, we note that the cognitive-experimental measures used in Cogsuite are intended to measure divergent processes typically not assessed in standard measures. Thus, for the set of cognitive-experimental measures, this analysis does not assume that standard measures represent a "gold-standard" for comparison as traditionally assumed in criterion validity analysis. Associations between Cogsuite and standard measures are reported for descriptive purposes and to identify associations between the two sets of measures that might suggest future research directions.

Attention—Attention Network Test

Condition effects (Fig. 2a,b). Each cue condition provides varying levels of information to predict target onset and location that should affect reaction time and accuracy. In order of lowest to highest accuracy and slowest to fastest reaction times, we predicted Invalid, None, Double and Valid cue types. As expected cue type was a significant predictor of reaction time $F(2.88, 1{,}020) = 22.30$, $P < 0.001$, with significant differences between every cue-pair (all pairwise comparisons <0.001). Valid cue reaction time was the fastest, followed by Double, None and Invalid cues. With regard to accuracy, no significant effect was observed for cue type. With regard to the target type (Congruent, Incongruent), we predicted that responses to Incongruent targets would be less accurate and slower than to Congruent targets. As predicted, a significant effect of target was observed for reaction time $F(1{,}354) = 31.45$, $P < 0.001$, with faster reaction times to Congruent targets than Incongruent targets. With regard to accuracy, a significant effect of target was observed $F(1{,}354) = 5.48$, $P = 0.02$, with better accuracy for Congruent targets than Incongruent targets, as expected.

Age and education effects. Age, but not education, was a significant predictor of reaction time $F(1{,}354) = 41.56$, $P < 0.001$, with increasing age leading to a slower reaction time, regardless of cue type. Education, but not age, was a significant predictor of accuracy $F(1{,}354) = 5.78$, $P = 0.017$, with higher education associated with higher accuracy, regardless of cue type.

Working Memory—N-back Test

Condition effects (Fig. 2c,d). Three levels of working memory load are presented with speed and accuracy expected to decrease with each level from 1- to 2- to 3-back. Working memory load (0,1,2,3-back) was marginally associated with reaction time $F(2.49{,}880.96) = 2.44$, $P = 0.06$, with significant differences between every load-pair, i.e., 1- versus 2-back, 1- versus 3-back, 2-versus 3-back (all pairwise comparisons <0.001). Reaction times increased with increasing load from 0- to 1- to 2-back, whereas 3-back reaction times were faster than 2-back reaction times. Load was not a significant predictor of accuracy, although significant differences were observed between all load-pairs (all pairwise comparisons <0.001) and accuracy was observed to decrease with increasing load.

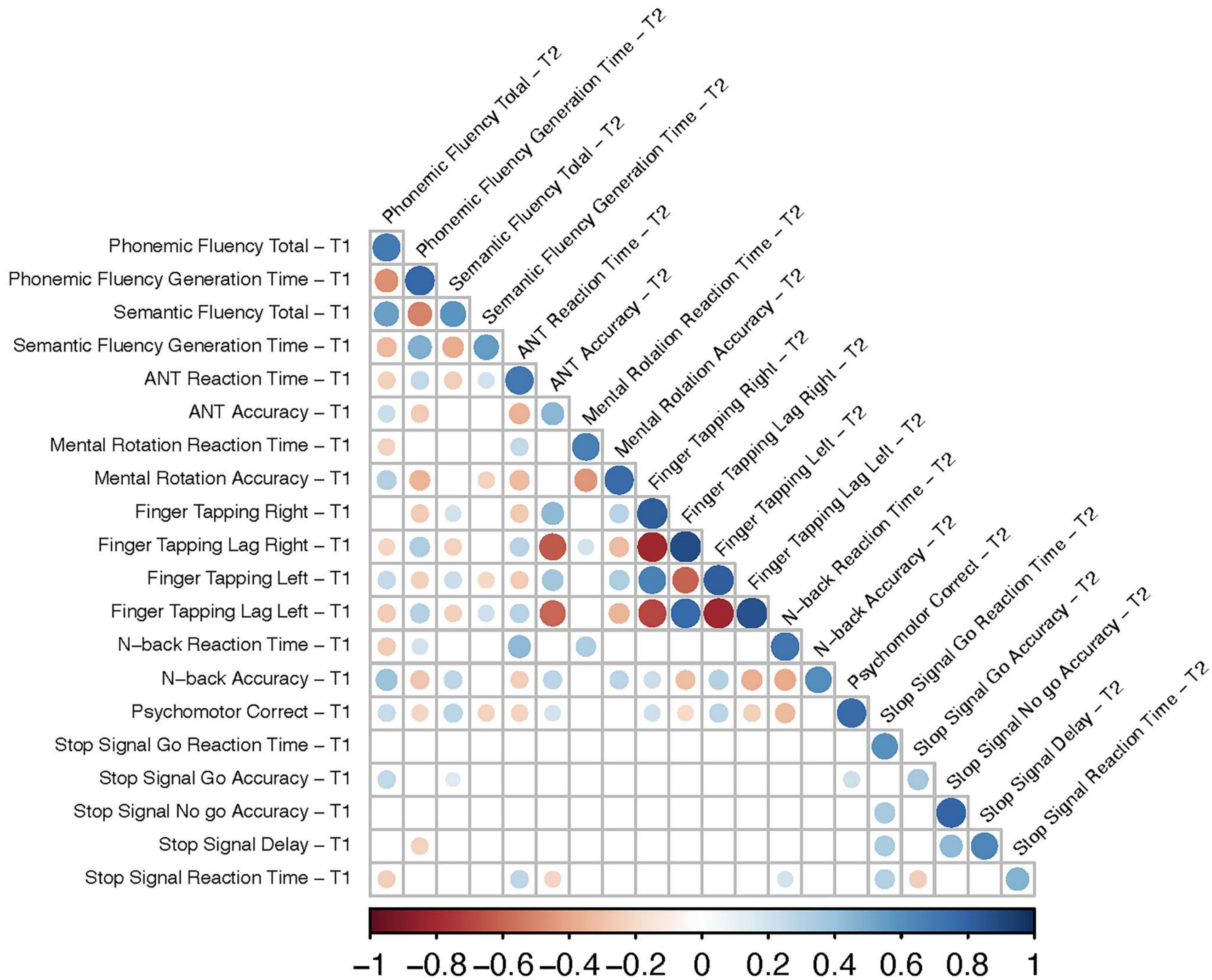

Fig. 1. Test-retest reliability of Cogsuite measures and summary intercorrelations of individual measures. All Pearson's r values are represented via the heatmap scale and range from −1 to 1, with warmer colors indicating positive correlations and cooler colors indicating negative correlations. The size and intensity of each circle indicates the strength of either negative or positive association. Empty cells indicate no significant correlation.

Age and education effects. Age, but not education, was a predictor of reaction time $F(1{,}354) = 28.63$, $P < 0.001$, with increasing age associated with slower reaction time. Age, but not education, was a predictor of accuracy $F(1{,}354) = 5.77$, $P = 0.017$, with increasing age associated witha lower accuracy.

Visuospatial Ability—Mental Rotation

Condition effects (Fig. 2e,f). Five conditions of rotation are presented with speed and accuracy expected to decrease with each increased degree of rotation. Degree of rotation was a significant predictor of reaction time $F(3.215,\ 1{,}138.132) = 9.101$, $P < 0.001$, with significant differences between every rotation pair (all pair-wise comparisons $P < 0.001$). Reaction times increased with increasing rotation load from 0° to 160° rotations. Degree of rotation was not a significant predictor of accuracy.

Age and education effects. Age and education were both significant predictors of reaction time $F(1{,}354) = 30.96$, $P < 0.001$ and $F(1{,}354) = 6.35$, $P = 0.012$, respectively, with increasing age and lower education associated with a slower reaction time. Age, but not education, was a significant predictor of accuracy $F(1{,}354) = 17.758$, $P < 0.001$, with increasing age leading to lower accuracy.

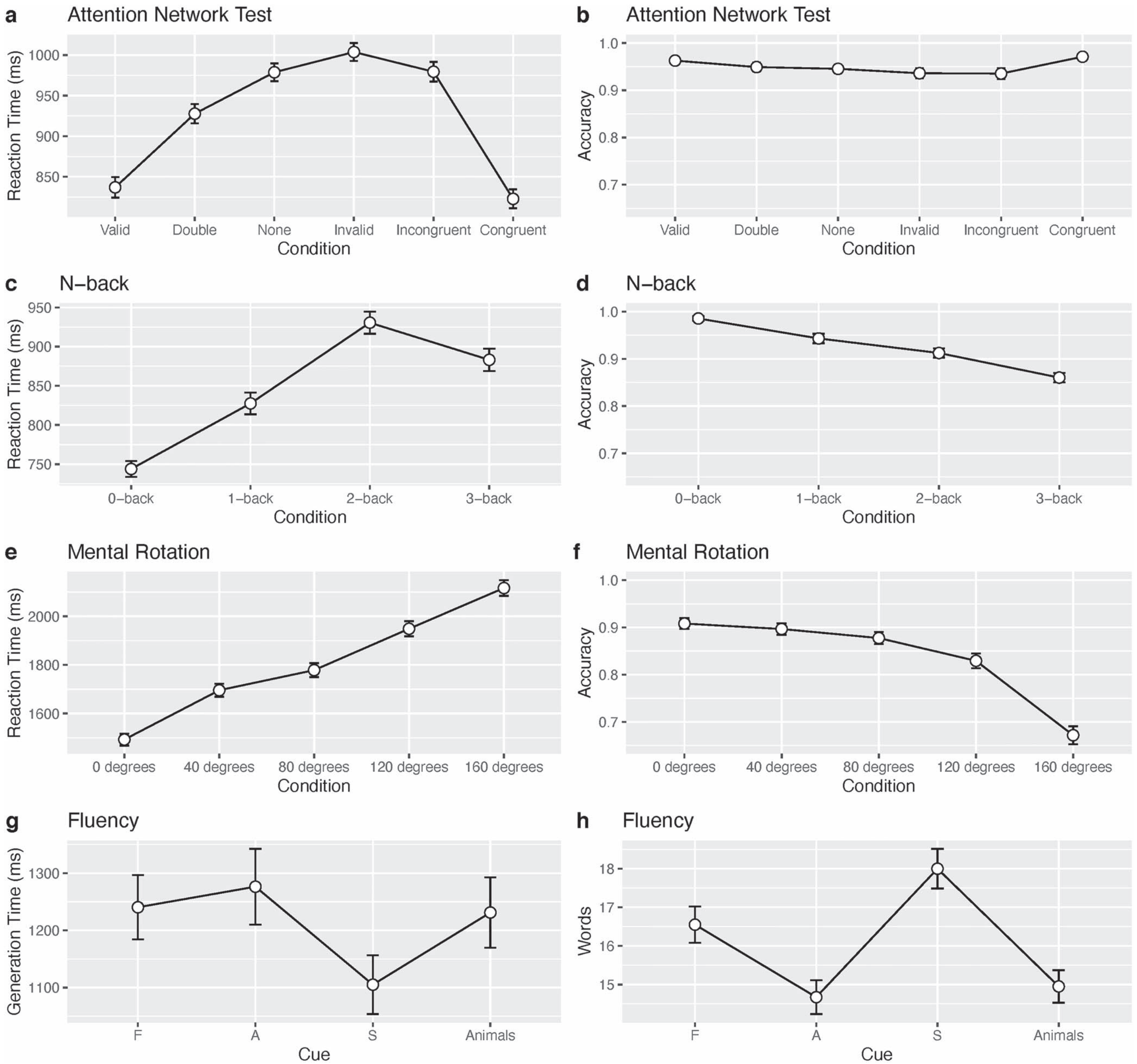

Fig. 2. Condition effects of Cogsuite measures.

Executive Functioning—Stop Signal Task

Age and education effects. For Go Trials, age, but not education, was a significant predictor of reaction time $F(1{,}357) = 9.69$, $P < 0.001$, with increasing age associated with a slower reaction time. For Mean Stop Signal Delay, age, but not education, was a significant predictor of stop signal intervals $F(1{,}346) = 4.86$, $P = 0.028$. No association was observed between either age or education and Stop Signal Reaction Time or No Go Trial accuracy.

Verbal Abilities—Verbal Fluency Task

Condition effects (Fig. 1g,h). Based on previous work reporting relative word rates for individual letters (Spreen & Strauss, 1991), we predicted a greater number of S words, followed by F and A words. A significant effect of letter was observed $F(1.981,$

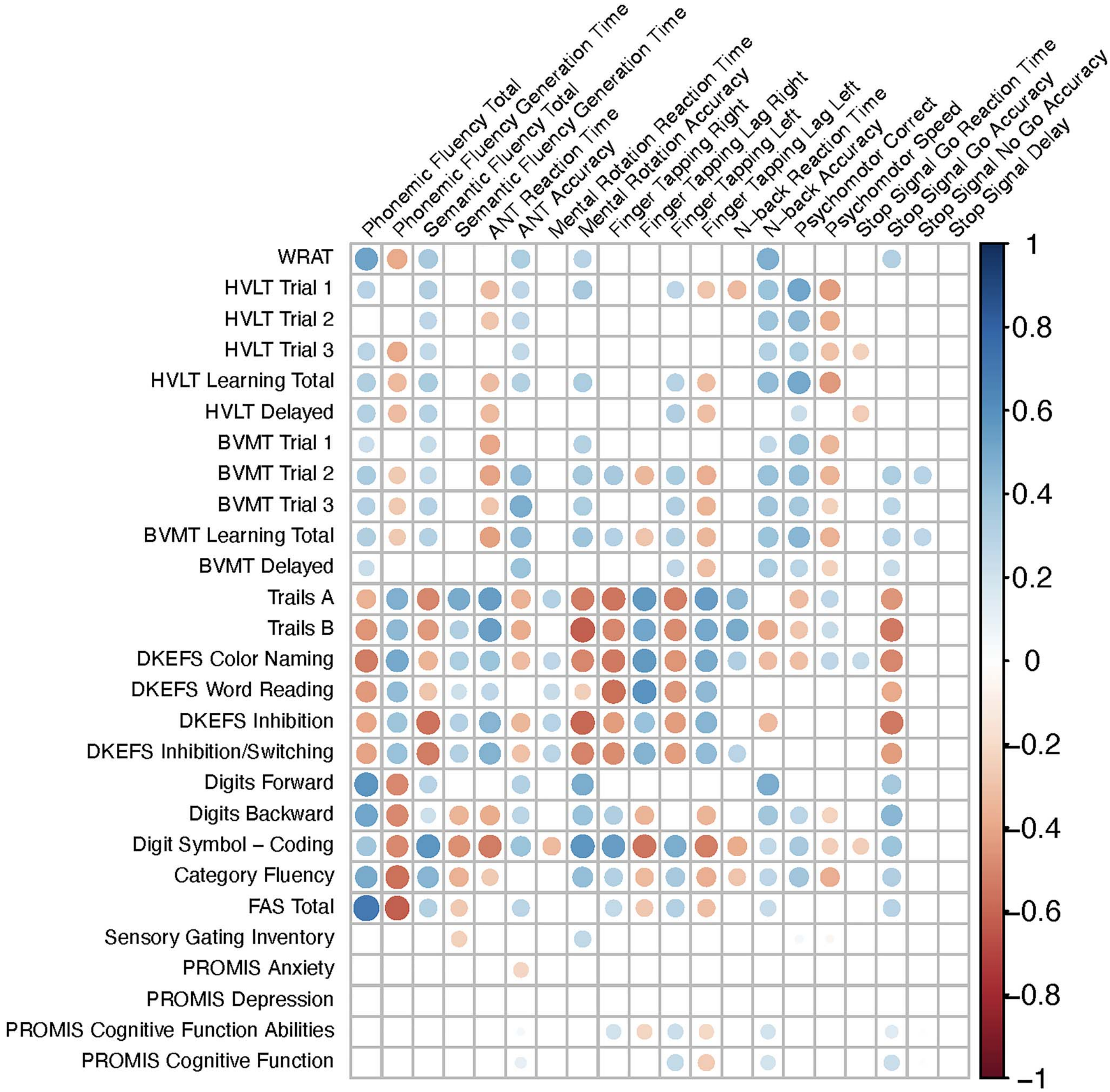

Fig. 3. Correlations between Cogsuite and standard measures. All Pearson's r values are represented via the heatmap scale and range from −1 to 1, with warmer colors indicating positive correlations and cooler colors indicating negative correlations. The size and intensity of each circle indicates the strength of either negative or positive association. Empty cells indicate no significant correlation.

701.285) = 6.36, $P < 0.002$; participants generated a greater number of S words, followed by F and A words, with significant differences between all letter-pairs (all pair-wise comparisons $P < 0.001$). With regard to phonemic fluency generation time, here defined as the time elapsed between the completion of one word and initiation of the next, a significant effect of letter was observed $F(2,706) = 3.38$, $P = 0.034$, with S words generated faster than F and A words.

Age and education effects. For the phonemic fluency task, age, but not education, was a significant predictor $F(1,354) = 5.59$, $P = 0.019$, with increasing age associated with fewer words. In the category fluency task (animals) age, but not education, was a significant predictor $F(1,354) = 12.41$, $P < 0.001$, with fewer words generated with increasing age.

Psychomotor Speed—Coding

Age and education effects. Age, but not education, was a significant predictor of coding reaction time $F(1,354) = 21.28$, $P < 0.001$ with increasing age leading to a slower reaction time. Similarly, age, but not education, was a significant predictor of total number of correct matches $F(1,354) = 18.09$, $P < 0.001$.

Motor Function—Finger Tapping

Condition effects. Right versus left hand of response was not a significant predictor of total taps on either side $F(1,349) = 2.66$, $P = 0.104$. A significant hand by dominant side (left dominant versus right dominant) interaction was observed $F(1,349) = 49.32$, $P < 0.001$, with right-side dominant individuals exhibiting fewer taps with their left hand than left-side dominant individuals.

Age and education effects. Age, but not education, was a significant predictor of total finger taps on both left and right sides $F(1,349) = 12.20$, $P < 0.001$.

Criterion Validity

A total of 76 datasets were available to assess the intercorrelations of Cogsuite variables with standard neuropsychological measures and self-report (Fig. 3). In general, stronger intercorrelations were observed between Cogsuite variables and measures of psychomotor speed, working memory and executive functioning, whereas verbal and visual learning measures exhibited weaker correlations with Cogsuite variables. The WRAT, used here as a measure of cognitive reserve, was most associated with verbal fluency measures and accuracy variables in choice reaction time tasks. With regard to self-reported depression and anxiety, only Attention Network Test Accuracy was weakly associated with anxiety and none were associated with depression. Regarding self-reported measures of cognition (Sensory Gating Inventory, Promis Cognitive Function, Promis Cognitive Abilities), generally weak correlations were observed in only a subset of variables.

Intercorrelations Between Cogsuite Measures

As expected, the pattern of intercorrelation of individual Cogsuite measures and variables generally groups into three levels of association (Supplementary Fig. S3). The strongest intercorrelations are for within task/between condition variables including in Verbal Fluency, Attention Network Test, N-back, Mental Rotation and Finger Tapping, indicating that the accuracy and speed of the response are consistent within a given cognitive measure regardless of the condition. The next strongest intercorrelations are observed between task/within metric variables, i.e., reaction time and accuracy, indicating that speed and/or accuracy are generally consistent across tasks regardless of cognitive processes tested. Finally, weaker but significant intercorrelations are observed between the reaction time and accuracy within tasks, with higher accuracy associated with a faster response, likely reflecting that the relative difficulty or ease of completion leads either to faster and more accurate responding or slower and less accurate response.

Discussion

The main goal in developing the Cogsuite battery was to assemble a series of cognitive-experimental and analog standard measures found to be sensitive to CRCD in past research to improve the resolution and sensitivity to cognitive dysfunction in survivorship. Here we have described psychometric characteristics of the battery, including reliability, internal validity (condition effects), external validity (known covariate effects in age and education) and criterion validity (the association with standard neuropsychological and self-reported measures). We believe the evidence presented supports the future use of Cogsuite as suitable measurement tool to supplement conventional CRCD assessments.

We have confirmed that all condition effects within each task are exhibited as predicted by previous literature. This indicates that visual and content changes to each task and the remote setting of administration did not alter the results so as to invalidate a given measure or variable. We found most measures and variables to exhibit moderate to strong reliability, except for variables in which range was restricted due to ceiling effects such as are found in measures of accuracy for relatively easy, choice reaction time tasks. Given that CRCD has been hypothesised to represent accelerated aging (Ahles et al., 2022), establishing sensitivity to aging effects on cognition was an important goal. We found significant associations of age with performance on every measure even in a relatively younger and age-restricted sample (ages 40–65). As standard measures are often affected by socio-economic variables such as education, we also sought to use measures that might be more invariant across education levels to reduce effects of SES on performance. We found that, in the majority of measures, education was not a significant predictor of performance with the exception of two variables in which a significant but small effect was found. With regard to criterion validity, i.e., correlations with standard measures, as we noted above, this form of validity typically assumes comparison with a "gold-standard" measure meant to assess the same behavior. In this case, the Cogsuite measures were chosen to better assess processes not captured by standard measures. Despite this difference, we were still able to find significant associations mainly in attention, psychomotor speed and executive function, and lesser associations with measures of learning and memory. With regard to associations between Cogsuite and self-reported measures, we found a similar pattern as that found in use of standard measures; self-reported and performance-based measures fail to be significantly associated in most cognitive syndromes including in CRCD (Bray, Dhillon, & Vardy, 2018). Finally, although a more modest sample was planned to be collected remotely, the SARS-CoV2 pandemic led us to collect most of the administrations remotely. As a result, the completion rates, condition effects, and test-retest data and results we report here reflect how the battery should behave when used as a remote battery for data collection.

One critically important comparison not reported in the current manuscript is whether Cogsuite can reliably detect differences between cancer survivors and non-cancer controls, which is the primary aim for the development of this battery. These group differences were reported in a separate manuscript (Gaynor et al., 2022b), primarily due to space constraints. In that manuscript, Cogsuite performance demonstrated sensitivity to cognitive differences between survivors and controls on multiple measures and variables, specifically in frontal-executive measures of attention and working memory. Taking these results together with the current analysis, Cogsuite appears to be reliable, valid and sensitive to CRCD, and yields greater resolution to specific cognitive processes that may be affected.

As with any research, our study is subject to limitations. As the majority of participants were collected remotely and were not treated at our institution, we relied on self-reported demographics, cancer history and other medical history, which may be less accurate than if verified through direct medical record review. Relatedly, although remote participants were instructed to take the Cogsuite battery in a quiet room free of distractions and without interruptions, remote administration is subject to less control over environmental variables, potential distractions during the administration of the battery, and participant effort, all of which could affect performance; we are reassured that all predicted condition effects in every measure were exhibited. Participant inclusion was subject to computer access and internet service, which may have tended to over-represent middle- to higher-SES individuals; we anticipate porting a subset of these measures to be administered on a tablet or smart-phone to improve access in the future. Likewise, we note that our sample is composed of generally highly educated women that may not be representative of national norms for education. Future work will focus on collecting data from a more educationally diverse sample to be reflective of the majority of individuals who may undergo assessment with this battery. Additionally, more breast cancer survivors were assessed in-office than non-cancer controls, resulting in an imbalance in the mode of administration that could not be corrected due to the onset of the pandemic, when in-office administration was halted. Additional data will be collected for in-office evaluations of non-cancer controls in the future. Finally, for most analyses, Pearson's correlation coefficient was used as the primary metric of validity and reliability. Pearson's r may not be as sophisticated as other metrics of predictive accuracy, such as area under the ROC curve, precision, recall, F-score and the confusion matrix. However, these methods all require that the ground truths are known with certainty, which is not available in the current study. Nevertheless, readers are reminded that the current effort is but the beginning of future work that aims to better capitalise on other methods to establish the accuracy of computer-based assessment tools.

The battery is anticipated to be used as an adjunct to standard neuropsychological assessment. Given the focus on cognitive experimental measures and methods in the development of this battery, key measures that are not included, e.g., learning and memory, can be administered as part of a core, standard battery in addition to the Cogsuite battery. The ability to administer these measures remotely and in an unsupervised setting will create less burden for survivors in participating in research on cognitive effects of CRCD and less burden on study personnel while, at the same time, preserving the fidelity, validity and sensitivity of the measures included. As a clinical tool, once appropriate normative data of a sufficient size are collected, Cogsuite has the potential to add additional information and resolution when used in addition to standard measures.

Supplementary Material

Supplementary material is available at *Archives of Clinical Neuropsychology* online.

Funding

This work was supported by grants from the National Cancer Institute at the National Institutes of Health (SBIR grant nos HHSN261201600024C, R01 CA218496-01A1, P30 CA008748).

Conflicts of Interest

Duane Jung is the CEO of Enformia, Inc., the publisher of the Cogsuite battery. All other authors confirm that they have no conflicts of interest.

Ethics approval

This study involving human participants was performed in accordance with the ethical standards of the institution and national research committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. The study protocol was approved by the ethics committee and the Institutional Review Boards at Memorial Sloan Kettering Cancer Center and City of Hope National Medical Center.

Consent to participate

Informed consent was obtained from all participants in this study.

Consent for publication

Not applicable.

Availability of data and materials

The data underlying this article will be shared on reasonable request to the corresponding author.

Code availability

Not applicable.

Author contributions

JCR: conceptualisation, funding acquisition, resources, investigation, data curation, statistical analysis, data visualisation; writing—original draft, review and editing; AG: data curation, formal analysis, writing—original draft, review and editing; AA: data curation, writing—review and editing; DJ: conceptualisation; funding acquisition, resources, investigation, data curation, writing- review and editing; LS: formal analysis, writing—review and editing; ER: writing—review and editing; YL: formal analysis, writing—review and editing; TAA: conceptualisation, investigation, writing—review and editing.

References

Ahles, T. A., & Root, J. C. (2018). Cognitive effects of cancer and cancer treatments. *Annual Review of Clinical Psychology*, *14*, 425–451. https://doi.org/10.1146/annurev-clinpsy-050817-084903.

Ahles, T. A., Saykin, A. J., McDonald, B. C., Li, Y., Furstenberg, C. T., Hanscom, B. S., et al. (2010). Longitudinal assessment of cognitive changes associated with adjuvant treatment for breast cancer: Impact of age and cognitive reserve. *Journal of Clinical Oncology*, *28*(*29*), 4434–4440. https://doi.org/10.1200/JCO.2009.27.0827.

Ahles, T. A., Schofield, E., Li, Y., Ryan, E., Root, J. C., Patel, S. K., et al. (2022). Relationship between cognitive functioning and frailty in older breast cancer survivors. *Journal of Geriatric Oncology*, *13*(*1*), 27–32. https://doi.org/10.1016/j.jgo.2021.07.011.

Barnett, J. H., Robbins, T. W., Leeson, V. C., Sahakian, B. J., Joyce, E. M., & Blackwell, A. D. (2010). Assessing cognitive function in clinical trials of schizophrenia. *Neuroscience and Biobehavioral Reviews*, *34*(*8*), 1161–1177. https://doi.org/10.1016/j.neubiorev.2010.01.012.

Army Individual Test Battery (1944). *Manual of directions and scoring*. Washington, DC: War Department, Adjunct General's Office.

Benedict, R. H., Schretlen, D., Groninger, L., Dobraski, M., & Shpritz, B. (1996). Revision of the Brief Visuospatial Memory Test: Studies of normal performance, reliability, and validity. *Psychological Assessment*, *8*(*2*), 145–153. https://doi.org/10.1037/1040-3590.8.2.145.

Boykoff, N., Moieni, M., & Subramanian, S. K. (2009). Confronting chemobrain: An in-depth look at survivors' reports of impact on work, social networks, and health care response. *Journal of Cancer Survivorship*, *3*(*4*), 223–232. https://doi.org/10.1007/s11764-009-0098-x.

Bray, V. J., Dhillon, H. M., & Vardy, J. L. (2018). Systematic review of self-reported cognitive function in cancer patients following chemotherapy treatment. *Journal of Cancer Survivorship*, *12*(*4*), 537–559. https://doi.org/10.1007/s11764-018-0692-x.

Chan, Y. H. (2003). Biostatistics 104: Correlational analysis. *Singapore Medical Journal*, *44*(*12*), 614–619. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/14770254.

Chao, H. H., Uchio, E., Zhang, S., Hu, S., Bednarski, S. R., Luo, X., et al. (2012). Effects of androgen deprivation on brain function in prostate cancer patients—a prospective observational cohort analysis. *BMC Cancer*, *12*(*1*), 371. https://doi.org/10.1186/1471-2407-12-371.

Chen, X., Li, J., Ren, J., Hu, X., Zhu, C., Tian, Y., et al. (2014). Selective impairment of attention networks in breast cancer patients receiving chemotherapy treatment. *Psychooncology*, *23*(*10*), 1165–1171. https://doi.org/10.1002/pon.3539.

Cherrier, M. M., Aubin, S., & Higano, C. S. (2009). Cognitive and mood changes in men undergoing intermittent combined androgen blockade for non-metastatic prostate cancer. *Psychooncology*, *18*(*3*), 237–247. https://doi.org/10.1002/pon.1401.

Conroy, S. K., McDonald, B. C., Smith, D. J., Moser, L. R., West, J. D., Kamendulis, L. M., et al. (2013). Alterations in brain structure and function in breast cancer survivors: Effect of post-chemotherapy interval and relation to oxidative DNA damage. *Breast Cancer Research and Treatment*, *137*(2), 493–502. https://doi.org/10.1007/s10549-012-2385-x.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan executive function system*. San Antonio, TX: Pearson.

Deprez, S., Amant, F., Yigit, R., Porke, K., Verhoeven, J., Van den Stock, J., et al. (2011). Chemotherapy-induced structural changes in cerebral white matter and its correlation with impaired cognitive functioning in breast cancer patients. *Human Brain Mapping*, *32*(*3*), 480–493. https://doi.org/10.1002/hbm.21033.

Deprez, S., Vandenbulcke, M., Peeters, R., Emsell, L., Smeets, A., Christiaens, M. R., et al. (2014). Longitudinal assessment of chemotherapy-induced alterations in brain activation during multitasking and its relation with cognitive complaints. *Journal of Clinical Oncology*, *32*(*19*), 2031–2038. https://doi.org/10.1200/JCO.2013.53.6219.

Dijkshoorn, A. B. C., van Stralen, H. E., Sloots, M., Schagen, S. B., Visser-Meily, J. M. A., & Schepers, V. P. M. (2021). Prevalence of cognitive impairment and change in patients with breast cancer: A systematic review of longitudinal studies. *Psychooncology*, *30*(*5*), 635–648. https://doi.org/10.1002/pon.5623.

Dyk, K. V., Crespi, C. M., Petersen, L., & Ganz, P. A. (2020). Identifying cancer-related cognitive impairment using the FACT-Cog perceived cognitive impairment. *JNCI Cancer Spectrum*, *4*(*1*), pkz099. https://doi.org/10.1093/jncics/pkz099.

Fan, J., Gu, X., Guise, K. G., Liu, X., Fossella, J., Wang, H., et al. (2009). Testing the behavioral interaction and integration of attentional networks. *Brain and Cognition*, *70*(*2*), 209–220. https://doi.org/10.1016/j.bandc.2009.02.002.

Ferguson, R. J., McDonald, B. C., Saykin, A. J., & Ahles, T. A. (2007). Brain structure and function differences in monozygotic twins: Possible effects of breast cancer chemotherapy. *Journal of Clinical Oncology*, *25*(*25*), 3866–3870. https://doi.org/10.1200/JCO.2007.10.8639.

Gaynor, A. M., Ahles, T. A., Ryan, E., Schofield, E., Li, Y., Patel, S. K., et al. (2022a). Initial encoding deficits with intact memory retention in older long-term breast cancer survivors. *Journal of Cancer Survivorship*, *16*(*5*), 940–947. https://doi.org/10.1007/s11764-021-01086-8.

Gaynor, A. M., Ahsan, A., Jung, D., Schofield, E., Li, Y., Ryan, E., et al. (2022b). Novel computerized neurocognitive test battery is sensitive to cancer-related cognitive deficits in survivors. *Journal of Cancer Survivorship*. https://doi.org/10.1007/s11764-022-01232-w.

Gevins, A. S., & Cutillo, B. C. (1993). Neuroelectric evidence for distributed processing in human working memory. *Electroencephalography and Clinical Neurophysiology*, *87*(*3*), 128–143. https://doi.org/10.1016/0013-4694(93)90119-G.

Hann, D. M., Jacobsen, P. B., Azzarello, L. M., Martin, S. C., Curran, S. L., Fields, K. K., et al. (1998). Measurement of fatigue in cancer patients: Development and validation of the fatigue symptom inventory. *Quality of Life Research*, *7*(*4*), 301–310. Retrieved from. https://doi.org/10.1023/A:1008842517972. http://www.ncbi.nlm.nih.gov/pubmed/9610214.

Hetrick, W. P., Erickson, M. A., & Smith, D. A. (2012). Phenomenological dimensions of sensory gating. *Schizophrenia Bulletin*, *38*(*1*), 178–191. https://doi.org/10.1093/schbul/sbq054.

Hooven, C. K., Chabris, C. F., Ellison, P. T., & Kosslyn, S. M. (2004). The relationship of male testosterone to components of mental rotation. *Neuropsychologia*, *42*(*6*), 782–790. https://doi.org/10.1016/j.neuropsychologia.2003.11.012.

Horowitz, T. S., Suls, J., & Trevino, M. (2018). A call for a neuroscience approach to cancer-related cognitive impairment. *Trends in Neurosciences*, *41*(*8*), 493–496. https://doi.org/10.1016/j.tins.2018.05.001.

Janelsins, M. C., Heckler, C. E., Peppone, L. J., Ahles, T. A., Mohile, S. G., Mustian, K. M., et al. (2018). Longitudinal trajectory and characterization of cancer-related cognitive impairment in a nationwide cohort study. *Journal of Clinical Oncology*JCO2018786624, *36*(*32*), 3231–3239. https://doi.org/10.1200/JCO.2018.78.6624.

Jim, H. S., Phillips, K. M., Chait, S., Faul, L. A., Popa, M. A., Lee, Y. H., et al. (2012). Meta-analysis of cognitive functioning in breast cancer survivors previously treated with standard-dose chemotherapy. *Journal of Clinical Oncology*, *30*(*29*), 3578–3587. https://doi.org/10.1200/JCO.2011.39.5640.

Kaufman, A. S., & Lichtenberger, E. O. (1999). *Essentials of WAIS-III assessment*. Hoboken, New Jersey Torrance, California: John Wiley & Sons Inc.

Kesler, S. R., Bennett, F. C., Mahaffey, M. L., & Spiegel, D. (2009). Regional brain activation during verbal declarative memory in metastatic breast cancer. *Clinical Cancer Research*, *15*(*21*), 6665–6673. https://doi.org/10.1158/1078-0432.CCR-09-1227.

Klaver, K. M., Duijts, S. F. A., Geusgens, C. A. V., Aarts, M. J. B., Ponds, R., van der Beek, A. J., et al. (2022). Neuropsychological test performance and self-reported cognitive functioning associated with work-related outcomes in occupationally active cancer survivors with cognitive complaints. *Journal of Cancer Survivorship*. https://doi.org/10.1007/s11764-022-01223-x.

Kramer, J. H., Mungas, D., Possin, K. L., Rankin, K. P., Boxer, A. L., Rosen, H. J., et al. (2014). NIH EXAMINER: Conceptualization and development of an executive function battery. *Journal of the International Neuropsychological Society*, *20*(*1*), 11–19. https://doi.org/10.1017/S1355617713001094.

Lezak, M. D. (1995). *Neuropsychological assessment*. Oxford, England: Oxford University Press.

Lipszyc, J., & Schachar, R. (2010). Inhibitory control and psychopathology: A meta-analysis of studies using the stop signal task. *Journal of the International Neuropsychological Society*, *16*(*6*), 1064–1076. https://doi.org/10.1017/S1355617710000895.

Logan, G. D., Cowan, W. B., & Davis, K. A. (1984). On the ability to inhibit simple and choice reaction time responses: A model and a method. *Journal of Experimental Psychology. Human Perception and Performance*, *10*(*2*), 276–291. https://doi.org/10.1037//0096-1523.10.2.276.

Luxton, J., Brinkman, T. M., Kimberg, C., Robison, L. L., Hudson, M. M., & Krull, K. R. (2014). Utility of the N-back task in survivors of childhood acute lymphoblastic leukemia. *Journal of Clinical and Experimental Neuropsychology*, *36*(*9*), 944–955. https://doi.org/10.1080/13803395.2014.957168.

Lyon, D. E., Cohen, R., Chen, H., Kelly, D. L., Starkweather, A., Ahn, H. C., et al. (2016). The relationship of cognitive performance to concurrent symptoms, cancer- and cancer-treatment-related variables in women with early-stage breast cancer: A 2-year longitudinal study. *Journal of Cancer Research and Clinical Oncology*, *142*(*7*), 1461–1474. https://doi.org/10.1007/s00432-016-2163-y.

Mandelblatt, J. S., Small, B. J., Luta, G., Hurria, A., Jim, H., McDonald, B. C., et al. (2018). Cancer-related cognitive outcomes among older breast cancer survivors in the thinking and living with cancer study. *Journal of Clinical Oncology*, *36*(*32*), 3211, JCO1800140–3222. https://doi.org/10.1200/JCO.18.00140.

McDonald, B. C., Conroy, S. K., Ahles, T. A., West, J. D., & Saykin, A. J. (2010). Gray matter reduction associated with systemic chemotherapy for breast cancer: A prospective MRI study. *Breast Cancer Research and Treatment*, *123*(*3*), 819–828. https://doi.org/10.1007/s10549-010-1088-4.

McDonald, B. C., Conroy, S. K., Ahles, T. A., West, J. D., & Saykin, A. J. (2012). Alterations in brain activation during working memory processing associated with breast cancer and treatment: A prospective functional magnetic resonance imaging study. *Journal of Clinical Oncology*, *30*(*20*), 2500–2508. https://doi.org/10.1200/JCO.2011.38.5674.

McDonald, B. C., Conroy, S. K., Smith, D. J., West, J. D., & Saykin, A. J. (2013). Frontal gray matter reduction after breast cancer chemotherapy and association with executive symptoms: A replication and extension study. *Brain Behavior and Immunity*, *30*(*Suppl*), S117–S125. https://doi.org/10.1016/j.bbi.2012.05.007.

Munir, F., Burrows, J., Yarker, J., Kalawsky, K., & Bains, M. (2010). Women's perceptions of chemotherapy-induced cognitive side affects on work ability: A focus group study. *Journal of Clinical Nursing*, *19*(*9–10*), 1362–1370. https://doi.org/10.1111/j.1365-2702.2009.03006.x.

Myers, J. S. (2010). Neuropsychologic testing for chemotherapy-related cognitive impairment. *Advances in Experimental Medicine and Biology*, *678*, 55–69. Retrieved from. http://www.ncbi.nlm.nih.gov/pubmed/20738007.

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385–392.

Robinson, L. M., Fitts, S. S., & Kraft, G. H. (1990). Laterality of performance in fingertapping rate and grip strength by hemisphere of stroke and gender. *Archives of Physical Medicine and Rehabilitation*, *71*(*9*), 695–698. Retrieved from. https://www.ncbi.nlm.nih.gov/pubmed/2375677.

Root, J. C., Andreotti, C., Tsu, L., Ellmore, T. M., & Ahles, T. A. (2016). Learning and memory performance in breast cancer survivors 2 to 6 years post-treatment: The role of encoding versus forgetting. *Journal of Cancer Survivorship*, *10*(*3*), 593–599. https://doi.org/10.1007/s11764-015-0505-4.

Root, J. C., Pergolizzi, D., Pan, H., Orlow, I., Passik, S. D., Silbersweig, D., et al. (2021). Prospective evaluation of functional brain activity and oxidative damage in breast cancer: Changes in task-induced deactivation during a working memory task. *Brain Imaging and Behavior*, *15*(*3*), 1364–1373. https://doi.org/10.1007/s11682-020-00335-1.

Root, J. C., Ryan, E., Barnett, G., Andreotti, C., Bolutayo, K., & Ahles, T. (2015). Learning and memory performance in a cohort of clinically referred breast cancer survivors: The role of attention versus forgetting in patient-reported memory complaints. *Psychooncology*, *24*(*5*), 548–555. https://doi.org/10.1002/pon.3615.

Rosano, C., Newman, A. B., Katz, R., Hirsch, C. H., & Kuller, L. H. (2008). Association between lower digit symbol substitution test score and slower gait and greater risk of mortality and of developing incident disability in well-functioning older adults. *Journal of the American Geriatrics Society*, *56*(*9*), 1618–1625. https://doi.org/10.1111/j.1532-5415.2008.01856.x.

Rosano, C., Perera, S., Inzitari, M., Newman, A. B., Longstreth, W. T., & Studenski, S. (2016). Digit symbol substitution test and future clinical and subclinical disorders of cognition, mobility and mood in older adults. *Age and Ageing*, *45*(*5*), 687–694. https://doi.org/10.1093/ageing/afw116.

de Ruiter, M. B., Reneman, L., Boogerd, W., Veltman, D. J., van Dam, F. S., Nederveen, A. J., et al. (2011). Cerebral hyporesponsiveness and cognitive impairment 10 years after chemotherapy for breast cancer. *Human Brain Mapping*, *32*(*8*), 1206–1219. https://doi.org/10.1002/hbm.21102.

Scherling, C., Collins, B., Mackenzie, J., Bielajew, C., & Smith, A. (2011). Pre-chemotherapy differences in visuospatial working memory in breast cancer patients compared to controls: An FMRI study. *Frontiers in Human Neuroscience*, *5*, 122. https://doi.org/10.3389/fnhum.2011.00122.

Scherling, C., Collins, B., Mackenzie, J., Bielajew, C., & Smith, A. (2012). Prechemotherapy differences in response inhibition in breast cancer patients compared to controls: A functional magnetic resonance imaging study. *Journal of Clinical and Experimental Neuropsychology*, *34*(*5*), 543–560. https://doi.org/10.1080/13803395.2012.666227.

Selamat, M. H., Loh, S. Y., Mackenzie, L., & Vardy, J. (2014). Chemobrain experienced by breast cancer survivors: A meta-ethnography study investigating research and care implications. *PLoS One*, *9*(*9*), e108002. https://doi.org/10.1371/journal.pone.0108002.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*(*3972*), 701–703. https://doi.org/10.1126/science.171.3972.701.

Simo, M., Gurtubay-Antolin, A., Vaquero, L., Bruna, J., & Rodriguez-Fornells, A. (2018). Performance monitoring in lung cancer patients pre- and post-chemotherapy using fine-grained electrophysiological measures. *NeuroImage: Clinical*, *18*, 86–96. https://doi.org/10.1016/j.nicl.2017.12.032.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the Stait-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Spreen, O., & Strauss, E. (1991). *A compendium of neuropsychological tests*. New York, NY: Oxford University Press.

Stern, R. A., & White, T. (2003). *Neuropsychological Assessment Battery*. WPS.

Suzumura, S., Kanada, Y., Osawa, A., Sugioka, J., Maeda, N., Nagahama, T., et al. (2021). Assessment of finger motor function that reflects the severity of cognitive function. *Fujita Medical Journal*, *7*(*4*), 122–129. https://doi.org/10.20407/fmj.2020-013.

Trevino, M., Zhu, X., Lu, Y. Y., Scheuer, L. S., Passell, E., Huang, G. C., et al. (2021). How do we measure attention? Using factor analysis to establish construct validity of neuropsychological tests. *Cognitive Research: Principles and Implications*, *6*(*1*), 51. https://doi.org/10.1186/s41235-021-00313-1.

Von Ah, D., Habermann, B., Carpenter, J. S., & Schneider, B. L. (2013). Impact of perceived cognitive impairment in breast cancer survivors. *European Journal of Oncology Nursing*, *17*(*2*), 236–241. https://doi.org/10.1016/j.ejon.2012.06.002.

Wefel, J. S., Lenzi, R., Theriault, R. L., Davis, R. N., & Meyers, C. A. (2004). The cognitive sequelae of standard-dose adjuvant chemotherapy in women with breast carcinoma: Results of a prospective, randomized, longitudinal trial. *Cancer*, *100*(*11*), 2292–2299. https://doi.org/10.1002/cncr.20272.

Wilkinson, G. S., & Robertson, G. J. (2006). *WRAT4: Wide Range Achievement Test*. Lutz, FL: Psychological Assessment Resources.

Zhai, F., Liu, J., Su, N., Han, F., Zhou, L., Ni, J., et al. (2020). Disrupted white matter integrity and network connectivity are related to poor motor performance. *Scientific Reports*, *10*(*1*), 18369. https://doi.org/10.1038/s41598-020-75617-1.