Asia-Pacific Journal of Oncology Nursing 10 (2023) 100309

Contents lists available at ScienceDirect

Asia-Pacific Journal of Oncology Nursing

journal homepage: www.apjon.org

Review

Assessing objective cognitive impairments in cancer survivors: Features and validity of measures for research and clinical applications

Sun Ok Jung [a], Jung Eun Esther Kim [b], Hee-Ju Kim [c,*]

[a] College of Nursing, Ewha Womans University, Seoul, Republic of Korea
[b] School of Nursing, San Francisco State University, San Francisco, USA
[c] College of Nursing, The Catholic University of Korea, Seoul, Republic of Korea

A R T I C L E I N F O

Keywords:
Cancer
Cognitive function
Cognitive impairment
Measure
Neuropsychological test

A B S T R A C T

Objective: This narrative review aims to (1) identify neuropsychological tests for assessing cognitive function impairment in patients with cancer, specifically in the domains of attention and memory, (2) summarize the characteristics of these tests, including cognitive function domains, test content, readability, and psychometric quality, and (3) evaluate the feasibility of each test in cancer care.

Methods: Data sources include published test manuals, documents from official web pages, and published journal articles.

Results: Our study identified eight neuropsychological tests that are most frequently used to assess the attention and memory domains of objective cognitive function in patients with breast cancer. These tests include the California Verbal Learning Test, Hopkins Verbal Learning Test, Rey Auditory Verbal Learning Test, Rey–Osterrieth Complex Figure, central nervous system (CNS) Vital Signs, Wechsler Adult Intelligence Scale, Wechsler Memory Scale, and Trail Making Test. They demonstrate acceptable evidence of psychometric quality and varying degrees of feasibility. Test feasibility is influenced by factors such as short testing time, brevity and comprehensiveness, clear cognitive domain distinctions, availability of normative data, minimal practice effects, ease of administration, and limited attention-span requirements. These attributes determine a test's feasibility for use in cancer care. Among the evaluated measures, the California Verbal Learning Test for memory, the Trail Making Test for attention, and the CNS Vital Signs for comprehensive assessment emerge as the most practical choices for cancer care.

Conclusions: The assessment and management of cognitive function impairment are crucial for enhancing the quality of life in cancer survivors. Nurses should possess knowledge of assessment tools for early detection and the ongoing monitoring of this symptom's progression.

Introduction

Patients with cancer report varying degrees of impairment in cognitive function during and after treatments.[1] Recent systematic reviews provided empirical evidence that impairments in cognitive function occur more prevalently and severely during or after initiating cancer treatment compared to baseline and healthy controls.[2–4] Patients report impairments in various areas of cognitive function, such as attention, memory, information processing speed, language, and executive function.[4–6]

Establishing practical and valid measures is fundamental to assessing cognitive-function impairments among patients with cancer. Although many researchers employ self-reported measures,[7] objectively assessing cognitive function is important in screening clinically meaningful impairments and investigating mechanisms of cognitive impairment. Researchers conducting objective assessments have employed a wide variety of neuropsychological tests originally developed for the evaluation of neuropsychological developments or disorders such as the Wechsler Adult Intelligence Scale,[8] Wechsler Memory Scale,[9] Trail Making Test,[10] central nervous system (CNS) Vital Signs,[11] and the Controlled Oral Word Association Test.[12] The International Cancer and Cognition Taskforce underscored that such variety in measures hampers evidence synthesis across studies and thus recommendes harmonizing measures.[13] However, systematic evaluations of these neuropsychological tests and guidance on how to select the tests best suited for patients with cancer are limited.

* *Corresponding author.*
E-mail address: heeju0906@gmail.com (H.-J. Kim).

https://doi.org/10.1016/j.apjon.2023.100309
Received 28 June 2023; Accepted 11 September 2023

Several points should be considered in selecting neuropsychological tests for patients with cancer, in addition to the evidence of psychometric quality. Patients with cancer are typically older adults, and they are easily fatigued and anxious during their clinical visits. Thus, tests should be easy and practical in application. Test scores should be adjusted for normal aging-associated changes in cognitive function.[14] Test scores should also be clinically interpretable. This is particularly important in screening subjects with significant cognitive impairments and in investigating the pathophysiological mechanisms underlying cognitive impairments. Moreover, impairment in cognitive function is mostly mild or moderate in patients with cancer,[6] necessitating a sensitive measure with a lower floor effect. The cognitive-function domain most impacted by cancer or its treatment is still uncertain.[15] Thus, researchers and clinicians need a domain-specific measure to investigate the impacted domain.

Last, when assessing the level of cognitive function repeatedly with a certain interval, examinees can memorize test items and perform better in the follow-up tests, the practice effect.[16] Practice effects can result in invalid conclusions on cognitive change over time and may have contributed to the inconsistency in findings across studies investigating the biobehavioral effects of cancer treatment on cognitive function.[1,17] Thus, tests with a small practice effect are preferred.

Each neuropsychological test has unique characteristics in testing format and in the focal cognitive domain, affecting its feasibility. Yet, no studies have provided comprehensive information on the characteristics of diverse neuropsychological tests to provide guidance for measurement selections to oncology researchers and clinicians. One review summarized the information on subjective and objective measures of cognitive impairment used in earlier studies (by the year 2011) for breast cancer survivors.[18] Yet, this previous review provided only brief information on the selected measures and did not include popularly used measures in recent longitudinal studies.

Therefore, the present study reviewed the measures frequently used for patients with breast cancer, the noncentral nervous system disease for which the issues of cognitive impairment were first raised. This population has been most studied within this research topic.[2,18,19] The impact of treatment on cognitive function has been more consistently reported in breast cancer than in other cancer types.[2] Female dominance and hormonal treatments in this disease type may increase the risk of the biochemical changes associated with cognitive function.[14]

Specific purposes of this narrative review were to (1) identify the neuropsychological tests assessing cognitive-function impairment for patients with breast cancer, (2) summarize the characteristics of such tests (eg, cognitive function domains, test contents, scoring, and psychometric quality), and (3) evaluate feasibility issues (eg, test duration, special needs, challenges, and practice effects) in applying to cancer patients.

Methods

This is a narrative review that identifies and synthesizes information on neuropsychological tests. We report this review based on a scale for the quality assessment of narrative review articles.[20]

Identification and selection of measures

Of note, this narrative review was planned as an extension of our previous systematic review, investigating the impact of chemotherapy on cognitive function in patients with breast cancer[4]; and thus used the data collected for the previous review to select measures.

For the present review, we identified and selected the measures in three steps (Fig. 1). In the first step, we identified the measures used to assess cognitive function in patients with breast cancer. This step was conducted solely based on data from our previous review. Our previous systematic review identified 42 studies and 81 measures across seven cognitive-function domains: attention, executive function, information-processing speed, verbal ability, motor function, visuospatial skill, and verbal/visual memory. The systematic approach to identify those studies and the measurement list by cognitive domains were reported in our previous review. We briefly summarize the search strategies and selection procedure of the previous review in Fig. 1 and Appendix 1. The present study started the selection procedure from this list of 81 measures. In this list, 21 tests were redundantly included in two or more cognitive-function domains, mostly subtests or subscores, not the test per se. This adjustment left 60 different tests (Table 1).

In the second step, we screened the measures and selected those used in three or more studies in each domain. As a wide variety of measures were used, we decided to focus on the frequently used measures and provide more comprehensive information for each test rather than briefly summarizing more measures. A total of 27 tests were used in three or more studies: six in the verbal/visual memory domain, five in the attention and executive-function domains, four in the information-processing-speed domain, three in the verbal-ability domain, and two in the motor-function and visuospatial-skills domains. Among them, 11 were included in two or more domains.

In the third step, we selected measures assessing the attention or memory domains. These domains were selected because they are required for many other cognitive functions, and the literature frequently reports them as impacted by cancer or its treatments.[4,6,18] A total of 11 tests met this inclusion criteria. Among them, two were listed in both domains, and two tests (Trail Making Test A and B) were subtests of one test. This leaves eight different tests for the present review: the California Verbal Learning Test,[21] Hopkins Verbal Learning Test,[22] Rey Auditory Verbal Learning Test,[23] Rey–Osterrieth Complex Figure,[24] the CNS Vital Signs,[11] Wechsler Adult Intelligence Scale,[8] Wechsler Memory Scale,[9] and the Trail Making Test.[10] Finally, we confirmed this final list included the measures recommended for the selected domains by the International Cancer and Cognition Taskforce.[13]

By applying this process, the scope of the review became reasonable and manageable. Two researchers (Drs. Jung and HJ Kim) performed the selection process. The two agreed at each step and jointly determined the inclusion and exclusion criteria.

Data collection for characteristics of the measures

Search strategy

A variety of databases was searched to retrieve publications for each test: Pubmed, EMbase, PsycINFO, and Google Scholar. Google was used to locate the manuals and official websites of each test. The data for each test were searched with sixteen key words and their relevant Mesh terms (Appendix 2): validity, construct validity, criterion validity, concurrent validity, performance validity, reliability, test-retest reliability, development, manual, normative, scoring, practice effect, ceiling, flooring, sensitivity, and cutoff. We limited our data search to English-language publications. Of note, we searched and selected the literature for each test purposefully. That is, the selection process cannot be presented like a systematic review.

Data extraction

Two researchers (Drs. Jung and JE Kim) retrieved and validated the information. When needed, the two researchers resolved differences through discussion with a third researcher (HJ Kim). The collected information included the published year, authors/copyright owners, target age, time to complete, contents, scoring, normative data, training for examiners, and reliability and validity evidence. We retrieved the highest validity and reliability evidence, if possible, from the most recent publication, from studies with adults aged 19 to 65, and from appropriate time interval for test-retest reliability. Although we did not limit the search to the English version of each test, most of the retrieved information was based on that version unless otherwise indicated.

Data analysis

Data syntheses

We set the following guidelines for synthesizing the data for each test. First, we determined the domain of each test, considering the labeling of test developers, labeling in previous reviews and studies, and finally this research team's opinion on the test content. Second, we considered the following suggested criteria in evaluating the psychometric quality of tests[25,26]: for internal consistency reliability, Cronbach's alpha between 0.70 and 0.95; for test-retest reliability, correlation between time points 0.60–0.90; and for criterion validity, correlation with criterion in statistical significance as well as in magnitude ($r > 0.30$ associated construct, $r > 0.50$ with similar construct). Of note, we set relatively liberal criteria for psychometric quality evaluation due to the diversity of study designs. For instance, the test-retest intervals were very long in many studies and diverse across subjects in a sample. Third, we examined criterion-validity evidence such that a test correlates well with other established neuropsychological tests. Yet, we did not differentiate between convergent and concurrent validity because authors used different names, and the gold standard for assessing the same domain of cognitive function was not well established.

Evaluating T feasibility

Feasibility evaluations were conducted based on the retrieved information and also through testing a trial version of a test, when available. Three researchers (Drs. Jung, JE Kim, and HJ Kim) independently evaluated each test's feasibility first. They then discussed these evaluations

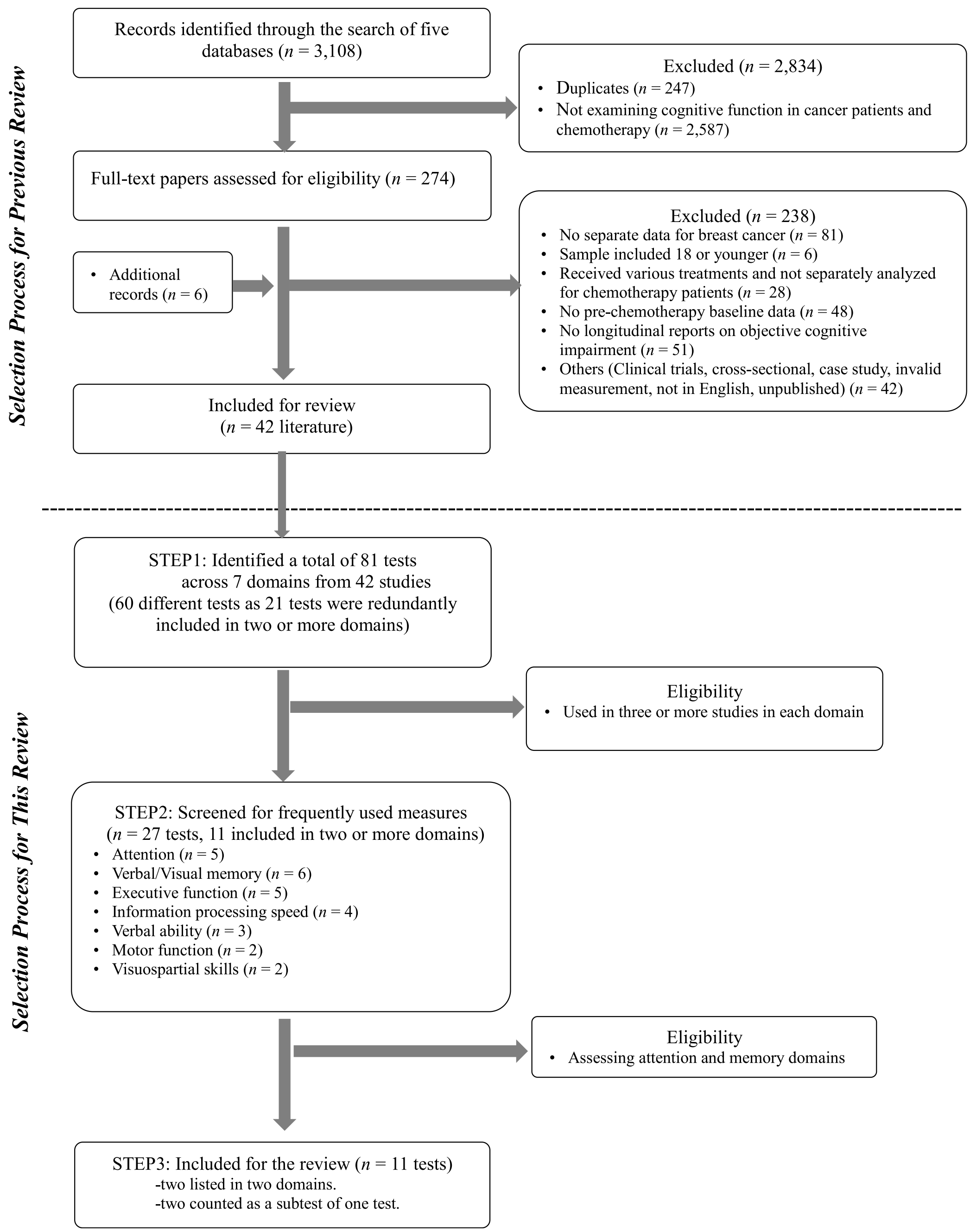

Fig. 1. Flow chart of measurement selection process.

Table 1
Neuropsychological tests identified from 42 studies.

No	Measures	No	Measures
1	Attentive matrices	31	Letter Cancellation Task
2	Auditory Consonant Trigrams Test	32	Line Tracing Test visual ability
3	Batterie d'efficiency mnésique	33	Montreal Cognitive Assessment*
4	Behavioral Assessment of the Dysexecutive Syndrome	34	National Institutes for Health Toolbox*
5	Benton Judgment of Line Orientation	35	Number Connection Test Type A
6	Benton Visual Retention Test	36	Paced Auditory Serial Addition Task
7	Boston Naming Test	37	Psychomotor Speed 9PEG
8	Brief Visuospatial Memory Test-Revised	38	Purdue Pegboard
9	Buschke Selective Reminding Task	39	Rapid Visual Processing Test
10	California Verbal Learning Test	40	Raven's Progressive Matrices
11	Cambridge Neuropsychological Test Automated Battery*	41	Regensburg Word Fluency Test*
12	Category fluency	42	Repeatable Battery for the Assessment of Neuropsychological Status*
13	Clock Drawing Test	43	Rey Auditory Verbal Learning*
14	CNS Vital Signs*	44	Rey–Osterrieth Complex Figure*
15	Cognitive Stability Index*	45	Rey Visual Learning Test
16	Continuous Performance Test*	46	Rivermead Story: recall
17	Controlled Oral Word Association Test*	47	Serial Dotting Test
18	D2 test	48	Short story
19	Delis–Kaplan Executive Function Scale*	49	Stroop test or Inter-Stroop
20	Digit Vigilance	50	Symbol Digit Modality Test*
21	Digital Symbol Test	51	Tests of Attentional Performance
22	Encoding Storage Retrieval	52	Trail Making Test A*
23	Eriksen Flanker Task	53	Trail Making Test B
24	Every day attention	54	Tower of London
25	Finger tapping	55	Verbal fluency
26	Grooved pegboard*	56	Verbal Learning and Memory Test
27	Grober and Buschke Test (G&B test)	57	Visual Reaction Time Test
28	General Practitioner Assessment of Cognition	58	Wechsler Abbreviated Scale of Intelligence*
29	High Sensitivity Cognitive Screen*	59	Wechsler Memory Scale*
30	Hopkins Verbal Learning Test	60	Wisconsin Card Sorting Test

CNS, central nervous system. This table lists 60 different tests alphabetically. The 21 tests included in two or more cognitive-function domains are marked with an asterisk. Our previous study reported a complete list of 81 tests classified by domains.

until a consensus was reached. We evaluated the strengths and weaknesses of each test in the following areas: testing duration, special needs, challenges, practice effects, and the influence of age, IQ, and education. We considered these issues while assuming we apply a test to patients with cancer in a research context as well as in a clinical context.

Results

Characteristics of eight selected tests

In Tables 2–5, we present the information for each test and our feasibility evaluations. The key characteristics of each test and our evaluations are summarized in the following subsections.

California Verbal Learning Test

The California Verbal Learning Test[21] evaluates verbal memory for subjects with a wide age range (16–90 years old; see Table 2). This simple test asks subjects to recall a given set of words. The test takes 30 minutes, with an additional 30 minuties for delayed recall. A 15-minute brief form is also available.[27] Normative data stratified by age, gender, education, and IQ are available for scoring.[28,29] Construct validity by factor analysis,[30] criterion validity with the Wechsler Adult Intelligence Scale,[28] and known-group validity to differentiate subjects with brain injury or neurological disorders[31,32] have been established. A good level of test-retest reliability (r = 0.80–0.84) was reported.[33] Considering the test characteristics, we concluded that it is easy to administer the test and to interpret the results for clinical use. However, the need for a silent setting can decrease its feasibility, and patients with hearing loss may be more challenged as the examiner presents the words verbally. As subjects might learn the list of words from a previous trial, the practice effect can confound the findings in repeated assessments with short intervals.

Table 2
California Verbal Learning Test & Hopkins Verbal Learning Test.

Measures	California Verbal Learning Test	Hopkins Verbal Learning Test
Target age	16–90	≥ 16
Domains	Verbal memory	Verbal memory
Assessment time (min)	For the standard tests: 30 + 30 (delay) For brief test: 15 + 15 (delay)	5–10
Contents	1. Asking to recall as many words as possible after one reading (five trials) (List A) 2. Asking to recall as many words as possible after one reading (List B) 3. Asking to recall the words from List A 4. Delayed tests after 20 min 5. Asking to remember the name of a set where the given word belongs	1. Part A (free recall): ask to repeat a list of words after verbally presenting a total of 12 words composed of four words from each of the three semantic categories (three trials) 2. Part B (Recognition): ask to recognize words from Part A after verbally presenting 24 words including 12 words from Part A (testing each word one by one)
Scoring	Correct answer	Correct answers
Normative data	Available by age, gender, education, IQ[28,29]	Available for four age-based groups[34]
Validity	1. Construct validity: six factors (general verbal learning, response discrimination, learning strategy, proactive effect, serial position effect, acquisition rate) discovered by factor analysis[30] 2. Criteria validity: r = 0.46 with the vocabulary subtest of WAIS[28] 3. Known group validity: significant group difference between traumatic brain injury/multiple sclerosis patients and normal controls[31,32]	1. Construct validity: three factors (memory, general cognitive processing, and recognition response bias) discovered by principal components analysis[35] 2. Criteria validity: significant correlations with the memory subtest of the WMS-R (r = 0.65–0.77), verbal IQ of the WAIS-R (r = 0.36–0.52), performance IQ of the WAIS-R (r = 0.40–0.49)[35] 3. Known group validity: significantly poorer performance in Alzheimer's Disease/vascular dementia than normal controls[35] 4. Predictive validity: 80% prediction accuracy to predict the disease group[35]
Reliability	1. Test-retest reliability: r = 0.80–0.84 by subtests, with a mean interval of 29 days[33]	1. Test-retest reliability: r = 0.42 (for false positive errors on recognition testing) – 0.66 by subscores with 9-month intervals[36]; 0.64 (for recognition discrimination index) – 0.81 by subtests with 6–8 week intervals[37]
Feasibility evaluation	Pros 1. Easy to administer, score and interpret 2. An alternate form is available Cons 1. Quiet room is required 2. Unsuitable for people with hearing loss 3. practice effect is expected.	Pros 1. Easy to administer, score and interpret 2. Short testing time 3. Six alternate evaluation forms are available Cons 1. Quiet room is required 2. Unsuitable for people with hearing loss 3. Severe practice effect

WMS-R, Wechsler Memory Scale Revised; WAIS, Wechsler Adult Intelligence Scale; WAIS-R, Wechsler Adult Intelligence Scale Revised.

Table 3
Rey Auditory Verbal Learning Test & Rey–Osterrieth Complex Figure.

Measures	Rey Auditory Verbal Learning Test	Rey–Osterrieth Complex Figure
Target age (year)	7–89	6–89
Domains	Verbal memory	Visual memory Executive function Motor function Visuospatial skill
Assessment time (min)	10–15 + 20 (delay)	45
Contents	1. Ask to recall as many words as possible after the evaluator reads a list of words (List A [15 words], five trials) 2. Ask to recall as many words as possible after the evaluator reads a new list of words (List B [15 words]) one time 3. Ask to recall words from List A without additional reading 4. Ask to recall words from List A 20 min later without an additional reading 5. Ask to find words from List A after showing 50 words including List A & B	1. Copy: ask to copy the given drawing figure 2. Immediate recall: ask to recall and draw the figure a min later 3. Delayed recall: ask to recall and draw the figure 20–30 min later
Scoring	Correct answers	Sum of 18 units scores Each unit scored from 0 to 2 depending on the level of completeness
Normative data	Normative data stratified by age, gender and education established[39]	Normative data stratified by age[47,48] Normative data stratified by age, gender and education established with French-speaking subjects[83]
Validity	1. Construct validity: three factors (memory, attention/learning, and inaccurate recall) discovered by the principal components analysis[40] 2. Criterion validity: $r = 0.68$ with WMS-R verbal index[41] 3. Known group validity: poorer performance among Alzheimer's disease and Parkinson's dementia patients than healthy controls[42] 4. Predictive validity: sensitivity 0.34–0.62 and specificity 0.10–0.99 to differentiate different levels of dementia, brain injury, and normal group[43]	1. Construct validity: one factor (visuospatial perception/memory) discovered by factor analysis[49] 2. Criterion validity: significant correlations with the memory subtest of the DTVP-A ($r = 0.33$–0.49), with the IQ of the Wechsler Intelligence scale ($r = 0.34$–0.39)[50], with dexterity and reaching motor tasks ($\beta = -0.35$)[51] 3. Known group validity: poorer performance in patients with Alzheimer's disease than normal controls[49]
Reliability	1. Test-retest reliability: $r = 0.68$ (for the sum of trials 1–5), with a 35-day interval[44] 2. Internal consistency: Cronbach's alpha, 0.80[44]	1. Test-retest reliability: $r = 0.73$ (for immediate recall), 0.79 (for delayed recall) with an average 251-day interval[54] 2. Internal consistency: Cronbach's alpha, 0.60 (for copy), 0.80 (for immediate recall), 0.82 (for delayed recall)[49] Split half, $r = 0.60$ (for copy), 0.84 (for immediate recall), 0.82 (for delayed recall)[49] 3. Interrater reliability: Kappa = 0.69–0.92 and Pearson $r = 0.80$–0.97[53]
Feasibility evaluation	Pros 1. Easy to administer 2. Short testing time 3. A supraspan memory test is possible due to the large stimulus load 4. Can test an unstructured learning strategy 5. Alternate testing formats are available Cons 1. Require a quiet room 2. Unsuitable for people with hearing loss 3. A practice effect is expected	Pros 1. Less influenced by language and culture 2. Can choose from a variety of versions to suit clinical and research needs Cons 1. Takes a long time 2. It is complicate to interpret test results 3. Need to establish agreement among raters 4. A practice effect is expected.

DTVP-A, Developmental Test of Visual Perception Adolescents and Adults; WMS-R, Wechsler Memory Scale Revised.

Hopkins Verbal Learning Test

The Hopkins Verbal Learning Test[22] is used to assess verbal memory for subjects aged 16 or higher (Table 2). It is a very simple test that asks subjects to recall a set of words and takes only 5–10 minutes. Normative data adjusted for age are available.[34] Construct validity by factor analysis, criterion validity with the Wechsler Adult Intelligence Scale and the Wechsler Memory Scale, and known-group and predictive validity to differentiate subjects with neurological disorders are well established.[35] Test-retest reliability[36,37] is acceptable for most subscores ($r > 0.60$). From our evaluations, its brevity and six testing-set options increase its feasibility for clinical use. However, because the test is verbally presented, it requires a silent setting, and patients with hearing loss may be more challenged. For longitudinal studies, practice effects are a concern. The lack of normative data adjusted for IQ or education is another weakness.

Rey Auditory Verbal Learning Test

The Rey Auditory Verbal Learning Test[23] is widely used in clinical settings for dementia screenings (Table 3). The test assesses verbal memory for subjects aged 7 to 89.[38] This simple test asks a subject to recall a set of words. It takes only 10–15 minutes, with an additional 20 minutes for delayed recall. Normative data adjusted for age, gender, and education are available.[39] Construct validity was confirmed by factor analysis.[40] Criterion validity[41] was established with the same construct. Known-group validity is established to differentiate a subject with neurological disorders.[42] Studies for predictive-validity evaluation have reported a wide range of sensitivity (0.34–0.62) and specificity (0.10–0.99).[43] Good internal consistency (alpha = 0.80) and acceptable test-retest reliability ($r = 0.68$) have been reported.[44] The short testing duration and assessments of diverse aspects of memory function are strengths of this test. The testing format can be adjusted.[45] The examiners need general education for psychological tests. The test also requires a silent setting, and patients with hearing loss may be more challenged.

Rey–Osterrieth Complex Figure

The Rey–Osterrieth Complex Figure[24] is a widely used neuropsychological test that assesses visual memory, executive function, motor function, and visuospatial skills (Table 3). This test can be applied to a wide age range (ages 6 to 89).[46] The Rey–Osterrieth Complex Figure Test is a drawing test in which respondents are asked to copy a presented figure. It takes 45 minutes to complete, including the delayed recall. For English-speaking subjects, normative data adjusted for age was established[47,48]; for French-speaking subjects, more diverse normative data were established. Construct validity was established by factor analysis.[49] The criterion validity was established by association with measures assessing memory[50] and motor skills.[51] Known-group validity was established by differentiating subjects with Alzheimer's disease.[49] In additional performance-validity evaluations, the test differentiated uncreditable performers well (specificity = 91%; sensitivity = 53%).[52] Good interrater reliability,[53] test-retest reliability,[54] and internal

Table 4
CNS Vital Signs & Wechsler Memory Scale.

Measures	CNS Vital Signs	Wechsler Memory Scale
Target age (year)	8–90	16–90
Domains	Attention Verbal memory Visual memory Information processing speed	Verbal memory Visual memory Attention Visuospatial skill
Assessment time (min)	45	60–90 25–35 for Abbreviated Scale
Contents	1. Verbal memory test: ask to recall 15 different words presented before another set of 15 words 2. Visual memory test: ask to recall 15 different geometric figures presented before another set of 15 figures 3. Finger Tapping Test: tap a space bar for 10 s with one hand's index finger (testing for each hand) 4. Symbol Digit Coding Test: ask to match symbols and digits that were presented as pairs previously 5. Stroop Test: ask to recognize the identity and difference between a colored word and a name of color 6. Shifting Attention Test: ask to match the color or shape of a presented figure with another figure following the instantly given instruction 7. Continuous Performance Test: ask to select the alphabet following the instruction	1. Logical memory: examiner presents two short stories and asks the examinee to tell the stories immediately and to recall the stories 30 min later 2. Verbal Paired Associates: examiner presents 10–14 word pairs and asks the examinee to recall the pair for the presented word 3. Visual Reproduction: ask the examinee to draw a picture after presenting a picture and to recall and draw it 15–30 min later 4. Design Memory: ask the examinee to relocate each design to the correct location on the grid after presenting the four to eight designs placed on the grid; retest after 15–30 min 5. Brief Cognitive Status Exam: screening for severe cognitive impairment with a brief test for the various cognitive function areas including time orientation, mental control, clock face drawing, recall, verbal fluency, etc. 6. Spatial Addition: ask to relocate blue dots to the right location on the grid after presenting two grids with red and blue dots 7. Symbol Span: ask to find the presented symbols in the right order after presenting a set of symbols
Scoring	Correct answer, errors, tapping number, time to complete a task	Correct answer
Normative data	Normative data stratified by age[11]	Normative data stratified based on age, gender, education level, race/ethnicity, and geographic regions within US[59]
Validity	1. Criterion validity: $r = 0.33$ (for memory domain score) – 0.56 (for visual memory) with logical memory of the WMS; $r = 0.26$ (for Shifting Attention Test errors) – 0.55 with computer-based neurobehavioral evaluation system[11] 2. Known group validity: significant difference between normative data and mild cognitive impairment/early dementia/post-concussion syndrome/severe traumatic brain injury/ADHD[11] 3. Predictive validity: sensitivity 0.12 (memory domain) – 0.35 (complex attention domain); specificity 0.88 (complex attention domain) – 0.99 (memory domain) in predicting neurological disorders[56]	1. Construct validity: two-dimensional structure (auditory learning/memory and visual attention/memory) established[60] 2. Criterion validity: $r = 0.40$ (for auditory index) −0.71 with WAIS index[59] 3. Known group validity: lower scores in people with epilepsy than the normative sample[61]
Reliability	1. Test-retest reliability: $r = 0.31$ (for Stroop test, errors) – 0.87 with the average 62-day interval[11]; $r = 0.11$ (for verbal memory) – 0.87, ICC = 0.10–0.86 between three separate sessions (each a week interval)[53]	1. Test-retest reliability: $r = 0.79$–0.82 for the five index scores with a 23-day interval; 0.59–0.76 for the seven subtests[62] 2. Internal consistency: Cronbach's alphas 0.93–0.96 for the five index scores; 0.82–0.97 for the seven subtest scores[62]
Feasibility evaluation	Pros 1. Comprehensive assessment 2. Easy to administer, and interpret 3. Easy to manage data and to construct database 4. Adjustable testing formats 5. Subtests can be separately used Cons 1. Requires computer settings in a quiet and independent place 2. Expensive 3. Lengthy to administer 4. Challenging for people with visual limitations or decreased motor function	Pros 1. Not boring with a play-like testing format 2. Easy to apply even to those who are anxious or have low concentration 3. Useful to assess memory based on attention 4. Logical memory subtests can be used separately Cons 1. Lengthy to administer 2. Unsuitable for individuals with speech impairments

CNS, central nervous system; ADHD, Attention Deficit Hyperactivity Disorder; WAIS, Wechsler Adult Intelligence Scale; WMS, Wechsler Memory Scale.

consistency[49] were reported (Table 3). As time intervals for test-retest are long, more evidence is needed to determine the test-retest reliability for the theoretically selected time interval. This test is less affected by language and culture, increasing its feasibility. Yet, testing time is relatively long. Examiners require experience or training for interrater agreement.[55]

CNS Vital Signs

CNS Vital Signs,[11] originally developed in computer format in 2006, assesses attention, verbal memory, visual memory, and information processing speed using seven subtests (Table 4). The target age range is wide, from 8 to 90.[11] The test takes 45 minutes for subjects to complete. Age-adjusted normative data are available.[11] The criterion validity for CNS Vital Signs has been wildly evaluated and shown to be acceptable.[11] Known-group validity shows the test can successfully detect the cognitive deficits associated with various neurological conditions, including traumatic brain injury and dementia.[11] However, additional evidence for predictive validity is needed because the test's sensitivity is low.[56] Test-retest reliability is acceptable for most subtests, but it is poor for some subtests.[57] Due to long time intervals or low reliability for some subtests, more evidence is needed to determine the test-retest reliability for the theoretically selected time interval. Because it is a computer-based test, it does not require the intensive involvement of examiners in testing and scoring. The testing domain and content can be customized for examinees,[58] decreasing practice effects in repeated tests. The database for test scores from examinees can be established.[11] Because the test is on a computer and takes longer to complete, older subjects or those with limited motor function or visual problems may be more challenged. An isolated setting with a computer is needed.

Wechsler Memory Scale

The Wechsler Memory Scale[9] assesses verbal and visual memory, attention, and visuospatial skills using seven subtests. The target age is 16–90.[59] The test takes 60–90 minutes to complete. A brief scale that only takes 25–35 minutes to complete is also available. Normative data adjusted for age, gender, education, and ethnicity are available. Evidence

Table 5
Wechsler Adult Intelligence Scale & Trail Making Test A, B.

Measures	Wechsler Adult Intelligence scale	Trail Making Test A, B
Target age (year)	16–90	18–89
Domains	Attention Information processing speed Memory Executive function	Attention Executive function Information processing speed Motor function
Assessment time (min)	60–90 15–30 for abbreviated scale	5–15
Contents	1. Verbal Comprehension Index a) Core Subtests - Similarity: ask the examinee to tell the similarity between words - Vocabulary: ask the examinee to tell the words of given objects and to define the given word - Information: ask the examinee about common senses b) Supplemental Subtests - Comprehension: ask the examinee about their understanding of given social situations or common concepts 2. Perceptual Reasoning Index a) Core Subtests - Block Design: build color blocks following the given pattern - Matrix Reasoning: find a proper image for the missing part in a drawing - Visual Puzzle: ask the examinee to choose three pieces to make up that pattern after showing a puzzle pattern in a stimulus book. b) Supplemental Subtests - Picture Completion: the examinee presents the missing part in the situation given by the drawing - Figure Weights: the examinee balances weights by choosing an object (figure) on the other side of the scale 3. Working Memory Index a) Core Subtests - Digit Span: ask the examinee to tell the numbers backward after memorizing the presented numbers - Arithmetic: ask the examinee to calculate mentally. b) Supplemental Subtests - Letter-Number Sequencing: ask the examinee to present the numbers in forward order and the alphabet in backward order 4. Processing Speed Index a) Core Subtests - Symbol Search: ask the examinee to find the instructed symbols - Coding: Time-limited task asking the examinee to match a digit with a symbol code using a key after showing a key in which the numbers 1 to 9 are each paired with a different symbol b) Supplemental Subtests - Cancellation: ask the examinee to cross out target pictures while looking at a random sequence of pictures	1. TMT A: ask examinee to draw a line between the consecutive numbers from 1 to 25 as fast as possible 2. TMT B: ask the examinee to draw a line between numbers (1–13) and the alphabets (A to L) in turn between categories and in consecutive order (1-A-2-B-3-C, etc.)
Scoring	Correct answers, total IQ score, index score based on subtests	Time to complete by second
Normative data	Normative data by age, gender, education, and race/ethnicity[63]	Normative data stratified by age, and education[68]
Validity	1. Construct validity: four factors (verbal comprehension, perceptual reasoning, working memory, and processing speed) discovered by factor analysis[63,64] 2. Criterion validity: correlated with RBANS total scores ($r = 0.75$)[63] 3. Divergent validity: verbal comprehension index of WAIS correlated with visual memory index of WMS III ($r = 0.20$–0.22)[63] 4. Known group validity: the difference between traumatic brain injury and normative database[65]	1. Construct validity: one factor (visual perception) established by factor analysis[69] 2. Criterion validity for Spanish: TMT A correlated with Stroop Color-Word ($r = -0.34$), WAIS-III Digit Symbol ($r = -0.63$), WAIS-III Digit Backward ($r = -0.50$), TMT B correlated with Stroop Color-Word ($r = -0.38$), WAIS-III Digit Symbol ($r = -0.57$), WAIS-III Digit Backward ($r = -0.54$)[70] 3. Known group validity: poorer performance in patients with Alzheimer's disease than normal controls[71]
Reliability	1. Test-retest coefficients: $r = 0.96$ for full scale IQ and 0.74–0.96 for primary indexes and subtests, with a mean interval of 22 days[63] 2. Inter-rater reliability: ICC = 0.91–0 0.97 across subtests[63] 3. Split-half reliability: $r = 0.78$–0.98 across age groups[63]	1. Test-retest reliability: $r = 0.7$ with a 235-day mean interval[54] 2. Inter-rater reliability: $r = 0.93$–0.99 for trail making test A; 0.88–0.99 for Trail Making Test B[72]
Feasibility evaluation	Pros 1. A comprehensive tool fpr assessing multiple domains 2. Subtests can be used separately Cons 1. More like an intelligence test 2. Require high concentration level 3. Lengthy to administer 4. Cannot be administered by nonpsychologists. 5. Not applicable for individuals with vision, auditory, or motor impairments 6. Very expensive	Pros 1. Easy to administer, and interpret 2. Short testing time 3. Require shorter attention span 4. Applicable in various settings 5. Different testing formats with colors are available Cons 1. Because memory is not assessed, the attention area may not be completely assessed 2. High ceiling effect is expected 3. May be affected by motor function 4. Severe practice effect

RBANS, The Repeatable Battery for the Assessment of Neuropsychological Status; WAIS, Wechsler Adult Intelligence Scale; WMS, Wechsler Memory Scale.

is well established for construct,[60] criterion[59] and known-group validity[61] (Table 4). Test-retest ($r \geq 0.59$) and internal consistency (Cronbach's alpha ≥ 0.82) are high.[62] The test is well structured and enjoyable because it is presented in a game format. Therefore, this test can be used with respondents who are anxious or who have a low attention span. Subtests can be used separately. However, its verbal memory subtest may not be suited for those with verbal language challenges.

Wechsler Adult Intelligence Scale

The Wechsler Adult Intelligence Scale[8] is a comprehensive test that assesses attention, information processing speed, memory, and executive function. Two or more subtests are included in each of the four subdomains. Subdomain index scores and total IQ scores can be calculated. The target age is from 16 to 90, and the test takes 60–90 minutes to complete.[63] An abbreviated scale takes 15–30 min to complete. Normative data adjusted for age, gender, education, and ethnicity are available.[63] The measurement construct is established by factor analysis.[63,64] Criterion validity has been well established by association with the established neurological testing battery.[63] The test differentiates subjects with traumatic brain injury well, showing acceptable known-group validity[65] (Table 5). All reliability indexes (test-retest, internal consistency, and interrater reliability) are high for all subtests and age groups.[63] Thus, this scale is often used as a criterion to evaluate the performance and embedded validity of other tests (ie, the credibility of test scores).[66] We concluded that the test provides a comprehensive evaluation of diverse cognitive functions, and the subtests are useful and practical. Yet, the core tests are similar to IQ tests and are highly affected by education and age. It is not applicable to those who are anxious, have a low attention span, or have vision, auditory, or motor impairments. The test should be given by trained psychologists for interpretation.[67]

Trail Making Test

The Trail Making Test[10] is a freely available neuropsychological test that assesses attention, executive function, information processing speed, and motor function. It is a simple drawing test targeted at those aged 18–89.[68] The test takes just 5–15 minutes to complete, and normative data for age and education are available.[68] Construct validity is established by factor analysis.[69] Criterion validity was established by association with measures assessing similar constructs (r ranging from -0.34 to -0.63).[70] This evidence is from nonEnglish-speaking subjects. Known-group validity[71] has been established, but predictive-validity evidence is limited (see Table 5). High interrater reliability (r ranged from 0.88–0.99)[72] and acceptable test-retest reliability ($r = 0.70$)[54] have been reported. Due to its brevity and simplicity in scoring and administering, it is highly feasible in busy clinical settings. The test requires a smaller attention span and, thus, can be applied to people who are anxious. Additionally, this test is less affected by aging. Yet, the test does not comprehensively assess attention because it does not evaluate memory. High practice effects can also be a concern. High-ceiling effects with low sensitivity can be expected for those with mild cognitive impairments.

Psychometric quality of tests

All tests provided reliability evidence for internal consistency and stability over time. Yet, several tests showed low test-retest reliability in some subtests (eg, Hopkins Verbal Learning Test[36,37] and CNS Vital Signs,[11]). Also, the time intervals between tests were too long (e.g., longer than a year) in many studies, which might have contributed to low test-retest reliability. Because certain cognitive-function domains can improve over time in younger participants and can naturally decrease in old age, it is important to select the theoretically relevant time window for the test-retest reliability of the neuropsychological tests. Thus, additional evidence is needed to confirm test-retest reliability for tests with low test-retest reliability using a theoretically relevant time window.

For the validity evaluation of the neuropsychological test, it is recommended to have an established relationship with the criterion or reference standard.[73] Criterion validity was evaluated for all tests. However, the selected criteria for most tests were not the gold standard for assessing the same cognitive domain. This indicates the need for more evidence for criterion validity with the gold standard.

As for the psychometric properties of the screening tests, predictive validity evidence is important because it shows how well a test can predict cognitive status. We considered known-group validity or sensitivity analyses as a predictive validity evaluation. We found known-group validity for all tests. A screening tool also requires good sensitivity and specificity,[73] but the evidence for sensitivity and specificity in most tests was either lacking or limited. Further, we did not find information on the ceiling or flooring effects of any tests.

Diagnostic accuracy for a few neuropsychological tests was also evaluated by performance validity, which reveals the effects of nonneurological factors (such as medication, indifference, pain, and fatigue) on testing scores.[75] In patients with cancer, the test results can also be confounded by anxiety. Yet, we did not find evidence for the performance validity of the measures in patients with cancer. Therefore, further psychometric evaluations as a screening tool are required for the target population.

Feasibility in application to patients with cancer

In Table 6, we summarized the key characteristics of the tests. All tests have been used and upgraded over decades, supporting their usefulness and quality. All tests are validated in various age groups and can be applied to adults up to 90 years old. Rey Auditory Verbal Learning Tests,[38] the Rey–Osterrieth Complex Figure,[76] and CNS Vital Signs,[11] can be applied to children under 12. The California Verbal Learning Test[77] and Wechsler Intelligence Scale[78] have children's versions.

Three tests (the California Verbal Learning Test, Hopkins Verbal Learning Test, and Rey Auditory Verbal Learning Test) mainly assess the verbal-memory domain and share the same feasibility issues. The strengths of these tests are the simplicity of the testing format and the short testing time. However, the tests require a quiet room to complete and can challenge those with hearing problems.

CNS Vital Signs, the Wechsler Memory Scale, and the Wechsler Adult Intelligence Scale are comprehensive batteries in that they consist of subtests and assess multiple cognitive functions, including attention and memory. Each test takes longer than 45 minutes to complete and requires a special testing setting. Thus, these tests may not be feasible in busy clinical settings. Yet, subtests of a comprehensive battery are more feasible. The Rey–Osterrieth Complex Figure and the Trail Making Test are not comprehensive batteries, but they do assess multiple cognitive functions. They are more feasible comprehensive tests considering their simplicity in administration and scoring.

The examiners are required to have a general education for psychological tests for all but CNS Vital Signs and the Trail Making Test. The Rey–Osterrieth Complex Figure and Wechsler Adult Intelligence Scale need additional training or experience for administration, scoring, and interpretation. Most tests can be administered in a paper-and-pencil setting as well as in a computer setting. However, CNS Vital Signs can only be administered in a computer setting. Computerized tests have several benefits, including no training requirements, lower scoring errors, shorter times, portability, and low cost.[18] Yet, computer settings may be challenging for the elderly, especially those over 70.

In Table 6, the issues surrounding each test's practice effects are presented. All tests have varying degrees of practice effects. CNS Vital Signs, the Wechsler Adult Intelligence Scale, and the Wechsler Memory Scale have lower practice effects.

Tests with normative data provide more clinically interpretable data. Considering that the performance on all tests is confounded by age and education to varying degrees and the patients with cancer are mostly older, normative data must be stratified at least by age. All tests established normative data by age. Most tests established normative data adjusted for education, except for the Hopkins Verbal Learning test and

Table 6
Characteristics of eight objective psycho-neurological tests.

Measures	Published year	Authors/copyright owner	Target age	Assessment time (minutes)	Domains (By the order of main focus)	Normative data		Special training for examiners	Online version	Cutoff to determine cognitive impairment	Practice effect	Language version
						Adjusted For education	Adjusted For age					
1. California Verbal Learning Test	1986 (ver.1) 2000 (ver.2) 2017 (ver.3)	Pearson	16–90	30 + 30 (delay) For brief test: 15 + 15 (delay)	Verbal memory	Yes	Yes	△	Web app via iPad for in-clinic/ remote testing	Yes	Moderate*	E, A, L,
2. Hopkins Verbal Learning Test	1991 1998 (revised)	PAR	≥ 16	5–10	Verbal memory	None	Yes	△	None	Yes	Severe*	E, A, L,
3. Rey Auditory Verbal Learning Test	1964	PAR	7–89	10–15 +20 (delay)	Verbal memory	Yes	Yes	△	None	Norm-based reading	Moderate*	E, A, L
4. Rey–Osterrieth Complex Figure	1944	PAR	6–89	45	Visual memory Executive function Motor function Visuospatial skill	Yes	Yes	○/△	None	Norm-based reading	Moderate*	E, A, L
5. CNS Vital Signs	2006	CNS Vital Signs	8–90	45	Attention Verbal memory Visual memory Information processing speed	None	Yes	Not necessary	Local, desktop app for in-clinic testing Web app for in-clinic/remote testing	Yes	Low	E, A
6. Wechsler Memory Scale	1945 (ver.1) 1987 (revised) 1997 (ver.3) 2009 (ver.4)	Pearson	16–90	60–90 For abbreviated scale: 25–35	Verbal memory Visual memory Attention Visuospatial skill	Yes	Yes	△	Web app via iPad for in-clinic/ remote testing	Norm-based reading	Low	E, A, L
7. Wechsler Adult Intelligence scale	1955 (ver.1) 1981 (ver.2) 1997 (ver.3) 2008 (ver.4) 2011 (abbr. ver.)	Pearson	16–90	60–90 For abbreviated scale: 15–30	Attention Information processing speed Memory Executive function	Yes	Yes	○/△	Web app via iPad for in-clinic/ remote testing	Norm-based reading	Low	E, A, L
8. Trail Making Test A, B	1944	Reitan Neuropsychology Laboratory (Publisher)	18–89	5–15	Attention Executive function Information processingspeed Motor function	Yes	Yes	Not necessary	Local, desktop app for in-clinic testing Purchased from Web app store	Yes	Severe*	E, A, L

○: Requiring training for the test. △: Requiring general education for psychological tests. *Practice effect may be manageable by using an alternate test format. Language versions were determined by the relevant validation studies. Cut-off data is only for English version. A: Asian language versions; E: English version; L: Latin-based language versions. CNS, central nervous system.

CNS Vital Signs. The Wechsler Memory Scale and Wechsler Adult Intelligence Scale have well-established normative data by age, gender, education, and race/ethnicity. The well-established cutoffs to determine the level of cognitive impairment are very useful for clinical and research contexts. The California Verbal Learning Test, Hopkins Verbal Learning Test, CNS Vital Signs, and the Trail Making Test provide useful cutoffs for screening subjects with significant cognitive impairment (dementia, mild cognitive impairment).

Age, gender, culture, and their interactions can influence performance on the tests.[79] The appropriateness of a test should be considered in a selected population. While the eight tests have been validated and used in diverse languages and cultures, they were developed in Western culture and the English language. The Rey Auditory Verbal Learning Test and Rey–Osterrieth Complex Figure were developed in France, though the latter is insensitive to language. CNS Vital Signs has 52 different language versions,[11,58] though their validation evidence in diverse languages is not well presented. The Trail Making Test has a version that is less sensitive to language. We noticed that some tests were modified to be appropriate to a certain culture or language and renamed (eg, Seoul Neuropsychological Screening Battery in Korea[80]). The relevance of those modified versions to patients with cancer should be re-evaluated in each culture.

Discussion

This narrative review identified eight neuropsychological tests frequently used in the literature investigating cognitive function in breast cancer and provided comprehensive information on those tests. It also discussed the practical issues of those tests in assessing cancer patients' cognitive function.

Overall evaluations and recommendations

We found acceptable evidence for psychometric quality for all of these tests. Yet, more validity evidence as a screening tool is needed, particularly for sensitivity and specificity, performance validity, and ceiling or flooring effects in the target population. The sensitivity and specificity of a test to detect mild levels of cognitive impairment in patients with cancer have been a special concern.[74] Future studies need to examine, at minimum, tests' flooring and ceiling effects for patients with cancer.

All tests are applicable to cancer patients, but each has unique strengths and weaknesses, yielding varying degrees of feasibility. To be feasible for patients with cancer, researchers have considered three points: a shorter testing time, taking less than 45 minutes to minimize fatigue,[81] brief tests that are comprehensive enough to detect an underlying problem,[73] and available demographically adjusted normative data.[73] We considered four additional points: smaller practice effects, clear distinction of assessed cognitive domains, ease of application, and lower attention-span requirements for use with anxious subjects. When using tests for repeated assessments, it is important to be aware of the potential for the practice effect. Yet, researchers can consider several strategies to minimize the impact of practice effects on test scores: increasing test-retest intervals; the use of dual baselines, alternate forms, and control groups; and statistical methods of calculating a reliable change index.[16] These can help ensure that changes in test scores over time are a true reflection of cognitive function rather than a result of practice effects.

Considering all of these consideration points, we concluded the Trail Making Test is most practical to assess attention, and the California Verbal Learning Test is best suited to assess verbal memory. These brief tests are handy for screening patients in a busy clinical setting. For a more comprehensive evaluation of both attention and memory and for more precise detection of cognitive function, CNS Vital Signs are very practical.

Implications and strengths

In the early years of the cancer-associated cognitive impairment study (2011), the International Cancer and Cognition Taskforce[13] suggested assessing the cognitive domains of learning, memory, and executive function and accordingly recommended measures for such domains: Hopkins Verbal Learning Test, Trail Making Test, and Controlled Oral Word Association. However, our recommendations differ. While the task-force recommendation is based on their expertise, our recommendation is based on the more recent and comprehensive literature on cancer-associated cognitive impairment and on neuropsychological tests. Studies conducted for the last two decades have used a wide variety of measures and frequently noted that the attention and memory domains are impaired in cancer patients.[4] Also, our recommendation is based on our feasibility evaluations for each test. We believe we provide useful and unique information for researchers and clinicians when selecting a measure.

Limitations

Nevertheless, we also acknowledge our limitations. First, we selected the tests based on studies in the breast-cancer literature and selected only those used frequently. We might have failed to identify other useful tests. Studies with other cancer types started later and followed the study designs of the breast-cancer literature, including measurement.[19] Also note that this review did not intend to introduce all possible measures, though that might be helpful. When those rarely used or newly developed tests become more visible in the literature, additional evaluations will be needed. Second, we acknowledge that we may have lost some information for each test despite our efforts to locate them. Lastly, we did not systematically evaluate issues of reliable change, performance validity, and cultural validation. Of note, this review did not evaluate the quality of tests in different cultures and languages. We also did not evaluate if test cutoffs are relevant in diverse cultural groups. These are important issues in measurement selection,[79,82] so they need more statistical evaluations in a separate study.

Conclusions

Assessing and managing symptoms is a critical role of oncology nurses. A body of evidence suggests that cancer patients, even those with disease or treatment not involving the central nervous system, experience impairment in cognitive function through diverse biological and behavioral mechanisms.[1,2,4] Thus, assessing and managing cognitive-function impairment is important to improve cancer survivors' quality of life. Nurses should be familiar with assessment tools for early detection and monitoring of this symptom. We recommend the California Verbal Learning Test to assess verbal memory, the Trail Making Test to assess attention, and CNS Vital Signs as a comprehensive assessment.

CRediT author statement

Hee-Ju Kim: Conceptualization, Methodology, Data curation, Formal analysis, Writing. **Sun Ok Jung:** Methodology, Data curation, Data collection, Writing – Revised draft preparation. **Jung Eun Esther Kim**: Formal analysis, Writing – Revised draft preparation, Data curation. All authors had full access to all the data in the study, and the corresponding author had final responsibility for the decision to submit for publication. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Declaration of competing interest

All authors have none to declare.

Funding

This study was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (Grant No. 2021R1A2C1094740). The funders had no role in considering the study

design or in the collection, analysis, interpretation of data, writing of the report, or decision to submit the article for publication.

Ethics statement

Not required.

Data availability statement

Data sharing not applicable – no new data generated.

Declaration of Generative AI and AI-assisted technologies in the writing process

No AI tools/services were used during the preparation of this work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apjon.2023.100309.

References

1. Ahles TA, Hurria A. New challenges in psycho-oncology research IV: cognition and cancer: conceptual and methodological issues and future directions. *Psycho Oncol*. 2018;27:3–9. https://doi.org/10.1002/pon.4564.
2. Kim HJ, Jung SO, Kim H, Abraham I. Systematic review of longitudinal studies on chemotherapy-associated subjective cognitive impairment in cancer patients. *Psycho Oncol*. 2020;29:617–631. https://doi.org/10.1002/pon.5339.
3. Kim HJ, Jung SO, Kim E, Abraham I. Association of chemotherapy and subjective cognitive impairment in breast cancer patients: meta-analysis of longitudinal prospective cohort studies. *Eur J Oncol Nurs*. 2022;57:102099. https://doi.org/10.1016/j.ejon.2022.102099.
4. Kim HJ, Kim JEE, Jung SO, Lee D, Abraham I. Neuropsychological effects of chemotherapy: systematic review of longitudinal studies on objective cognitive impairment in breast cancer patients. *Cancer Nurs*. 2023;46:E159–E168. https://doi.org/10.1097/NCC.0000000000001079.
5. Von Ah D, Habermann B, Carpenter JS, Schneidger BL. Impact of perceived cognitive impairment in breast cancer survivors. *Eur J Oncol Nurs*. 2013;17:236–241. https://doi.org/10.1016/j.ejon.2012.06.002.
6. Bernstein LJ, McCreath GA, Komeylian Z, Rich JB. Cognitive impairment in breast cancer survivors treated with chemotherapy depends on control group type and cognitive domains assessed: a multilevel meta-analysis. *Neurosci Biobehav Rev*. 2017; 83:417–428.
7. Bray VJ, Dhillon HM, Vardy JL. Systematic review of self-reported cognitive function in cancer patients following chemotherapy treatment. *J Cancer Surviv*. 2018;12: 537–559. https://doi.org/10.1007/s11764-018-0692-x.
8. Wechsler D. *Wechsler adult intelligence scale*. New York: The Psychological Corporation; 1955.
9. Wechsler D. *Wechsler Memory Scale*. 1945.
10. Reitan RM. *Army Individual Test Battery. Manual of Directions and Scoring*. Washington, DC: War Department, Adjutant General's Office; 1944.
11. Gualtieri CT, Johnson LG. Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs. *Arch Clin Neuropsychol*. 2006;21:623–643. https://doi.org/10.1016/j.6.05.0acn.20007.
12. Benton AL, Hamsherde SK, Sivan AB. *Multilingual Aplasia Examination*. 2nd ed. Iowa City, IA: AJA Associates; 1983.
13. Wefel JS, Vardy J, Ahles T, Schagen SB. International Cognition and Cancer Task Force recommendations to harmonize studies of cognitive function in patients with cancer. *Lancet Oncol*. 2011;12:703–708. https://doi.org/10.1016/S1470-2045(10)70294-1.
14. Joly F, Giffard B, Rigal O, et al. Impact of cancer and its treatments on cognitive function: advances in research from the Paris international cognition and cancer task force symposium and update since 2012. *J Pain Symptom Manag*. 2015;50(6): 830–841. https://doi.org/10.1016/j.jpainsymman.2015.06.019.
15. Ono M, Ogilvie JM, Wilson JS, et al. A meta-analysis of cognitive impairment and decline associated with adjuvant chemotherapy in women with breast cancer. *Front Oncol*. 2015;5:59, 10.3389%2Ffonc.2015.00059.
16. Calamia M, Markon K, Tranel D. Scoring higher the second time around: meta-analyses of practice effects in neuropsychological assessment. *Clin Neuropsychol*. 2012;26(4):543–570. https://doi.org/10.1080/13854046.2012.680913.
17. Li M, Caeyenberghs K. Longitudinal assessment of chemotherapy-induced changes in brain and cognitive functioning: a systematic review. *Neurosci Biobehav Rev*. 2018;92: 304–317. https://doi.org/10.1016/j.neubiorev.2018.05.019.
18. Cheung YT, Tan EH, Chan A. An evaluation on the neuropsychological tests used in the assessment of post chemotherapy cognitive changes in breast cancer survivors. *Support Care Cancer*. 2012;20:1361–1375. https://doi.org/10.1007/s00520-012-1445-4.
19. Di Iulio F, Cravello L, Shofany J, et al. Neuropsychological disorders in non-central nervous system cancer: a review of objective cognitive impairment, depression, and related rehabilitation options. *Neurol Sci*. 2019;40(9):1759–1774. https://doi.org/10.1007/s10072-019-03898-0.
20. Baethge C, Goldbeck-Wood S, Mertens S. SANRA—a scale for the quality assessment of narrative review articles. *Research integrity and peer review*. 2019;4(1):1–7. https://doi.org/10.1186/s41073-019-0064-8.
21. Delis DC, Kramer JH, Kaplan E, Ober BA. *California Verbal Learning Test Research Edition Manual*. San Antonio: The Psychological Corporation; 1987.
22. Brandt J. The Hopkins Verbal Learning Test: development of a new memory test with six equivalent forms. *Clin Neuropsychol*. 1991;5(2):125–142.
23. Rey A. *L'examen clinique en psychologie*. Paris, France: Presses Universitaires de France; 1964.
24. Rey A. L'examen psychologique dans les cas d'encéphalopathie traumatique. (Les problems.). *Arch Psychol*. 1941;29:286–340.
25. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60: 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012.
26. Nunnally JC, Bernstein IH. *Psychometric Theory*. 3rd ed. New York, NY: McGraw-Hill; 1994.
27. Delis DC, Kramer JH, Kaplan E, Ober BA. *CVLT3 California Verbal Learning Test Third Edition Manual*. San Antonio, TX: The Psychological Corporation; 2017.
28. Delis DC, Kramer JH, Kaplan E, Ober BA. *California Verbal Learning Test*. 2nd ed. San Antonio, TX: The Psychological Corporation; 2000. Adult Version. Manual.
29. Wiens AN, Tindall AG, Crossen JR. California Verbal Learning Test: a normative data study. *Clin Neuropsychol*. 1994;8:75–90. https://doi.org/10.1080/01688639708403853.
30. Delis DC, Freeland J, Kramer JH, Kaplan E. Integrating clinical assessment with cognitive neuroscience: construct validation of the California Verbal Learning Test. *J Consult Clin Psychol*. 1988;56:123–130. https://doi.org/10.1037/0022-006X.56.1.123.
31. Jacobs ML, Donders J. Criterion validity of the California Verbal Learning Test-(CVLT-II) after traumatic brain injury. *Arch Clin Neuropsychol*. 2007;22:143–149. https://doi.org/10.1016/j.acn.2006.12.002.
32. Stegen S, Stepanov I, Cookfair D, et al. Validity of the California verbal learning test–II in multiple sclerosis. *Clin Neuropsychol*. 2010;24:189–202. https://doi.org/10.1080/13854040903266910.
33. Woods SP, Delis DC, Scott JC, Kramer JH, Holdnack JA. The California Verbal Learning Test–second edition: test–retest reliability, practice effects, and reliable change indices for the standard and alternate forms. *Arch Clin Neuropsychol*. 2006;21: 413–420. https://doi.org/10.1016/j.acn.2006.06.002.
34. Benedict RH, Schretlen D, Groninger L, Brandt J. Hopkins Verbal Learning Test–Revised: normative data and analysis of inter-form and test–retest reliability. *Clin Neuropsychol*. 1998;12:43–55.
35. Shapiro AM, Benedict RH, Schretlen D, Brandt J. Construct and concurrent validity of the Hopkins verbal learning test–revised. *Clin Neuropsychol*. 1999;13:348–358.
36. Rasmusson DX, Bylsma FW, Brandt J. Stability of performance on the Hopkins verbal learning test. *Arch Clin Neuropsychol*. 1995;10:21–26.
37. O'Neil-Pirozzi TM, Goldstein R, Strangman GE, Glenn MB. Test–re-test reliability of the Hopkins Verbal Learning Test-Revised in individuals with traumatic brain injury. *Brain Inj*. 2012;26:1425–1430. https://doi.org/10.3109/02699052.2012.694561.
38. Schmidt M. *Rey Auditory Verbal Learning Test: A Handbook*. vol. 17. Los Angeles, CA: Western Psychological Services; 1996.
39. Loring DW, Saurman JL, John SE, Bowden SC, Lah JJ, Goldstein FC. The rey auditory verbal learning test: cross-validation of mayo normative studies (MNS) demographically corrected norms with confidence interval estimates. *J Int Neuropsychol Soc*. 2023;29(4):397–405. https://doi.org/10.1017/S1355617722000248.
40. Weitzner DS, Pugh EA, Calamia M, Roye S. Examining the factor structure of the Rey auditory verbal learning test in individuals across the life span. *J Clin Exp Neuropsychol*. 2020;42(4):406–414. https://doi.org/10.1080/13803395.2020.1741517.
41. Callahan CD, Johnstone B. The clinical utility of the Rey Auditory-Verbal Learning Test in medical rehabilitation. *J Clin Psychol Med Settings*. 1994;1:261–268. https://doi.org/10.1007/BF01989627, 1994.
42. Tierney MC, Nores A, Snow WG, Fisher RH, Zorzitto ML, Reid DW. Use of the rey auditory verbal learning test in differentiating normal aging from alzheimer's and Parkinson's dementia. *Psychol Assess*. 1994;6(2):129–134. https://doi.org/10.1037/1040-3590.6.2.129.
43. Poreh A, Tolfo S, Krivenko A, Teaford M. Base-rate data and norms for the rey auditory verbal learning embedded performance validity indicator. *Appl Neuropsychol: Adult*. 2017;24:540–547.
44. de Sousa Magalhães S, Malloy-Diniz LF, Hamdan AC. Validity convergent and reliability test–retest of the rey auditory verbal learning test. *Clin Neuropsychiatry*. 2012;9:129–137.
45. Hawkins KA, Dean D, Pearlson GD. Alternative forms of the Rey auditory verbal learning test: a review. *Behav Neurol*. 2004;15(3–4):99–107.
46. Meyers J, Meyers KR. *Rey Complex Figure Test and Recognition Trial*. Psychological Assessment Resources; 1995.
47. Fastenau PS, Denburg NL, Hufford BJ. Adult norms for the rey-osterrieth complex figure test and for supplemental recognition and matching trials from the extended complex figure test. *Clin Neuropsychol*. 1999;13(1):30–47. https://doi.org/10.1076/clin.13.1.30.1976.
48. Machulda MM, Ivnik RJ, Smith GE, et al. Mayo's older Americans normative studies: visual form discrimination and copy trial of the rey–osterrieth complex figure. *J Clin Exp Neuropsychol*. 2007;29(4):377–384. https://doi.org/10.1080/13803390600726803.

49. Berry DT, Allen RS, Schmitt FA. Rey–Osterrieth complex figure: psychometric characteristics in a geriatric sample. *Clin Neuropsychol*. 1991;5:143–153. https://doi.org/10.1080/13854049108403298.
50. Smith SR, Chang J, Schnoebelen KJ, Edwards JW, Servesko AM, Walker SJ. Psychometrics of a simple method for scoring organizational approach on the Rey–Osterrieth Complex Figure. *J Neuropsychol*. 2007;1:39–51. https://doi.org/10.1348/174866407x180800.
51. Lingo J, VanGilder, Lohse KR, Duff K, Wang P, Schaefer SS. Evidence for associations between Rey–Osterrieth Complex Figure test and motor skill learning in older adults. *Acta Psychol*. 2021;214:103261.
52. Sugarman MA, Holcomb EM, Axelrod BN, Meyers JE, Liethen PC. Embedded measures of performance validity in the Rey Complex Figure test in a clinical sample of veterans. *Appl Neuropsychol: Adult*. 2016;23:105–114. https://doi.org/10.1080/23279095.2015.1014557.
53. Deckersbach T, Savage CR, Henin A, et al. Reliability and validity of a scoring system for measuring organizational approach in the Complex Figure Test. *J Clin Exp Neuropsychol*. 2000;22:640–648. https://doi.org/10.1076/1380-3395 (200010)22:5;1-9;ft640.
54. Levine AJ, Miller EN, Becker JT, Selnes OA, Cohen BA. Normative data for determining significance of test–retest differences on eight common neuropsychological instruments. *Clin Neuropsychol*. 2004;18:373–384. https://doi.org/10.1080/1385404049052420.
55. Shin MS, Park SY, Park SR, et al. Clinical and empirical applications of the rey–osterrieth complex figure test. *Nat Protoc*. 2006;1(2):892–899. https://doi.org/10.1038/nprot.2006.115.
56. Brooks BL, Sherman EM, Iverson GL. Embedded validity indicators on CNS Vital Signs in youth with neurological diagnoses. *Arch Clin Neuropsychol*. 2014;29:422–431. https://doi.org/10.1093/arclin/acu029.
57. Littleton AC, Register-Mihalik JK, Guskiewicz KM. Test–retest reliability of a computerized concussion test: CNS Vital Signs. *Sport Health*. 2015;7:443–447. https://doi.org/10.1177/1941738115586997.
58. *CNS Vital Signs*. Academic Research; 2022. https://www.cnsvs.com/AcademicResearch.html/. Accessed August 1, 2022.
59. Maccow G. WMS-IV: administration, scoring, basic interpretation. http://images.pearsonclinical.com/images/Products/WMS-IV/WMS-IV_Webinar_September_2011_Handout.pdf. Accessed April 20, 2023.
60. Hoelzle JB, Nelson NW, Smith CA. Comparison of Wechsler Memory Scale–fourth edition (WMS–IV) and third edition (WMS–III) dimensional structures: improved ability to evaluate auditory and visual constructs. *J Clin Exp Neuropsychol*. 2011;33:283–291. https://doi.org/10.1080/13803395.2010.511603.
61. Moore PM, Baker GA. Validation of the Wechsler Memory Scale-Revised in a sample of people with intractable temporal lobe epilepsy. *Epilepsia*. 1996;37:1215–1220. https://doi.org/10.1076/jcen.20.2.280.1161.
62. Drozdick LW, Holdnack JA, Hilsabeck RC. *Essentials of WMS-IV Assessment*. vol. 85. Hoboken, NJ: John Wiley & Sons; 2011.
63. Wechsler D. Wechsler adult intelligence scale. In: *Technical and Interpretive Manual*. 4th ed. San Antonio, TX: Pearson; 2008.
64. Nelson JM, Canivez GL, Watkins MW. Structural and incremental validity of the Wechsler Adult Intelligence Scale–fourth edition with a clinical sample. *Psychol Assess*. 2013;25:618–630.
65. Carlozzi NE, Kirsch NL, Kisala PA, Tulsky DS. An examination of the Wechsler Adult Intelligence Scales, (WAIS-IV) in individuals with complicated mild, moderate and severe traumatic brain injury (TBI). *Clin Neuropsychol*. 2015;29:21–37. https://doi.org/10.1080/13854046.2015.1005677.
66. Erdodi LA, Abeare CA. Stronger together: the Wechsler Adult Intelligence Scale—fourth Edition as a multivariate performance validity test in patients with traumatic brain injury. *Arch Clin Neuropsychol*. 2020;35:188–204. https://doi.org/10.1093/arclin/acz032.
67. Ryan JJ, Schnakenberg-Ott SD. Scoring reliability on the Wechsler adult intelligence scale-(WAIS-III). *Assessment*. 2003;10(2):151–159. https://doi.org/10.1177/1073191103010002006.
68. Tombaugh TN. Trail Making Test A and B: normative data stratified by age and education. *Arch Clin Neuropsychol*. 2004;19:203–214.
69. Groff MG, Hubble LM. A factor analytic investigation of the Trail Making Test. *Clin Neuropsychol*. 1981;3:11–13.
70. Sánchez-Cubillo I, Periáñez JA, Adrover-Roig D, et al. Construct validity of the Trail Making Test: role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *J Int Neuropsychol Soc*. 2009;15:438–450. https://doi.org/10.1017/s1355617709090626.
71. Amieva H, Lafont S, Auriacombe S, et al. Analysis of error types in the Trail Making Test evidences an inhibitory deficit in dementia of the Alzheimer type. *J Clin Exp Neuropsychol*. 1998;20:280–285.
72. Poreh A, Miller A, Dines P, Levin J. Decomposition of the Trail Making Test-reliability and validity of a computer assisted method for data collection. *Arch Assess Psychol*. 2012;2:57–72.
73. Vardy J, Rourke S, Tannock IF. Evaluation of cognitive function associated with chemotherapy: a review of published studies and recommendations for future research. *J Clin Oncol*. 2007;25:2455–2463. https://doi.org/10.1200/jco.2006.08.1604.
74. Lange M, Castel H, Le Fel J, et al. How to assess and manage cognitive impairment induced by treatments of non-central nervous system cancer. *Neurosci Biobehav Rev*. 2019;107:602–614. https://doi.org/10.1016/j.neubiorev.2019.09.028.
75. Greher MR, Wodushek TR. Performance validity testing in neuropsychology: scientific basis and clinical application—a brief review. *J Psychiatr Pract*. 2017;23:134–140. https://doi.org/10.1097/pra.0000000000000218.
76. Conson M, Siciliano M, Baiano C, et al. Normative data of the rey–osterrieth complex figure for Italian-speaking elementary school children. *Neurol Sci*. 2019;40:2045–2050. https://doi.org/10.1007/s10072-019-03929-w.
77. Talley JL. *Children's Auditory Verbal Learning Test-2. Professional Manual*. Odessa, FL: Psychological Assessment Resources Inc; 1993.
78. Wechsler D. *Wechsler Intelligence Scale for Children*. 5th ed. Bloomington, MN: Pearson; 2014.
79. Makri E, Giannouli V. Cross-cultural cognitive and affective differences in aging: can culture shape the expression and perception of psychopathology in old age? *Encephalos*. 2022;59:34–43. http://www.encephalos.gr/pdf/59-4-01e.pdf.
80. Kang Y, Na DL, Hahn SJ. *Seoul Neuropsychological Screening Battery*. Incheon: Human Brain Research & Consulting; 2003.
81. Hurria A, Lachs M. Is cognitive dysfunction a complication of adjuvant chemotherapy in the older patient with breast cancer? *Breast Cancer Res Treat*. 2007;103:259–268. https://doi.org/10.1007/s10549-006-9383-9.
82. Andreotti C, Root JC, Schagen SB, et al. Reliable change in neuropsychological assessment of breast cancer survivors. *Psycho Oncol*. 2016;25(1):43–50. https://doi.org/10.1002/pon.3799.
83. Tremblay MP, Potvin O, Callahan BL, et al. Normative data for the Rey–Osterrieth and the Taylor complex figure tests in Quebec-French people. *Arch Clin Neuropsychol*. 2015;30:78–87. https://doi.org/10.1093/arclin/acu069.