

COMMENTARY

Understanding the Profile of Cancer-Related Cognitive Impairments: A Critique of Meta-Analyses

Todd S. Horowitz, Melissa Treviño, Ingrid M. Gooch, Korrina A. Duffy

See the Notes section for the full list of authors' affiliations.

Correspondence to: Todd S. Horowitz, PhD, National Cancer Institute, 9609 Medical Center Dr, Rockville, MD 20850 (e-mail: todd.horowitz@nih.gov).

Abstract

A large body of evidence indicates that cancer survivors who have undergone chemotherapy have cognitive impairments. Substantial disagreement exists regarding which cognitive domains are impaired in this population. We suggest that is in part due to inconsistency in how neuropsychological tests are assigned to cognitive domains. The purpose of this paper is to critically analyze the meta-analytic literature on cancer-related cognitive impairments (CRCI) to quantify this inconsistency. We identified all neuropsychological tests reported in seven meta-analyses of the CRCI literature. Although effect sizes were generally negative (indicating impairment), every domain was declared to be impaired in at least one meta-analysis and unimpaired in at least one other meta-analysis. We plotted summary effect sizes from all the meta-analyses and quantified disagreement by computing the observed and ideal distributions of the one-way χ^2 statistic. The actual χ^2 distributions were noticeably more peaked and shifted to the left than the ideal distributions, indicating substantial disagreement among the meta-analyses in how neuropsychological tests were categorized to domains. A better understanding of the profile of impairments in CRCI is essential for developing effective remediation methods. To accomplish this goal, the research field needs to promote better agreement on how to measure specific cognitive functions.

Cancer treatment is a major success story of modern medicine. As cancer mortality rates decrease, the number of cancer survivors has been steadily increasing and is projected to grow from 15.5 million to 20 million over the next decade in the United States (1). Accordingly, concerns about the problems faced by cancer survivors are growing. Among these concerns are cognitive impairments, which have been linked to cancer treatments, particularly chemotherapy, leading to colloquial descriptions such as “chemobrain” or “chemofog.” Researchers, however, have adopted the theory-neutral term cancer-related cognitive impairment (CRCI).

A substantial body of research over the last two decades indicates that CRCI is a real and persistent problem (2). However, it is difficult to identify exactly what the problem is. For example, patients' subjective reports of cognitive problems often do not correlate well with objective neuropsychological tests (3–6); this problem is not unique to CRCI (see for example [7,8]). Additionally, each investigator seems to have a different definition of what counts as impairment, leading to widely

differing estimates of the prevalence of CRCI (9). Furthermore, it is not yet clear to what extent these cognitive impairments are caused by cancer treatments, as opposed to the disease itself and the multifaceted stress of becoming a “cancer patient” (10,11).

Whatever the source of CRCI, what can be said about which cognitive functions are impaired? Deficits have been reported across a gamut of neuropsychological domains, including executive function, verbal working memory, visuospatial functions, processing speed, reaction time, and attention (12,13). The problem is that there is little consistency across findings. Different research groups have come to different conclusions about which domains are impaired and which are spared. There are two primary reasons for this confusion. First, researchers disagree on which neuropsychological tests tap which cognitive functions, making it difficult to compare results from different studies (internal consistency). Second, the assignment of neuropsychological tests to cognitive domains in neuropsychology does not always match up with the way cognitive psychologists

Received: September 28, 2018; Revised: April 19, 2019; Accepted: May 16, 2019

Published by Oxford University Press 2019. This work is written by US Government employees and is in the public domain in the US.

assess and conceptualize cognition, making it difficult to interpret findings and make connections across disciplines (*external consistency*). In this paper, we address the internal consistency issue, specifically with respect to meta-analyses.

Knowing which cognitive functions are likely to be impaired, and which are not, is valuable for three main reasons. Identifying which specific cognitive functions are impaired will 1) provide patients who are about to undergo cancer therapy with accurate information on what side effects to expect from their course of treatment, 2) allow researchers to develop better countermeasures to improve patients' functioning, and 3) help us pin down the affected neural networks in the brain, which in turn will help us understand the physiological pathways underlying CRCI. By providing quantitative summaries of the empirical literature, meta-analyses can give us a clearer picture of the research findings on CRCI. However, differences in the classification of neuropsychological tests across meta-analyses contribute to the inconsistency of our picture of CRCI.

Summary of Meta-Analyses

Research on the cognitive sequelae of cancer treatments has been conducted since the early 1970s (14,15). In the last four decades, a plethora of conventional narrative reviews have been published (eg, [16,17]). The first meta-analysis on the topic was published by Anderson-Hanley et al. in 2003 (18). Meta-analysis is a statistical technique for combining information from multiple independent studies to achieve better estimates of effect sizes. The purpose of meta-analysis is to supplement narrative reviews of the scientific literature with a more systematic methodology. A meta-analysis involves extracting effect size estimates from each study, weighting them according to their variance (ie, studies with lower variance, typically those with larger sample sizes, are weighted more heavily than those with higher variance), and then combining them to get a summary effect size, which is a more reliable estimate of the true effect than any individual study could provide. Instead of reasoning by counting up studies on one side of an issue or the other (eg, eight studies say X, whereas only two studies say not-X, therefore X is probably true), one can report an effect size estimate and the accompanying confidence intervals (eg, the effect of X was 2.3, and the 95% confidence intervals did not include zero). This ability to make precise, principled statements about the state of the evidence should therefore bring clarity to a scientific field.

We searched for meta-analyses in the Scopus database using the search terms ("chemotherapy" and "meta-analysis" and "cognitive" or "cognition"). We restricted the results to analyses of objective neuropsychological testing in survivors of nonpediatric, noncentral nervous system cancers. We decided not to include one paper because it was designed to measure the sensitivity of the tests, rather than the nature of CRCI (19). This search yielded seven papers (18,20-25) (Note that an eighth meta-analysis [26] was published after the analyses herein were conducted.) In addition, we unaccountably omitted a ninth meta-analysis by Lindner et al. (27). All seven papers provide evidence for cognitive deficits in cancer patients who have undergone chemotherapy, but disagree on which specific cognitive functions are impaired.

In this paper, we will use the term function to refer to a theoretical cognitive ability that has psychological reality, whereas the term *domain* will be used to refer to an artificial category or label for a set of behavioral tests. Ideally, domains correspond

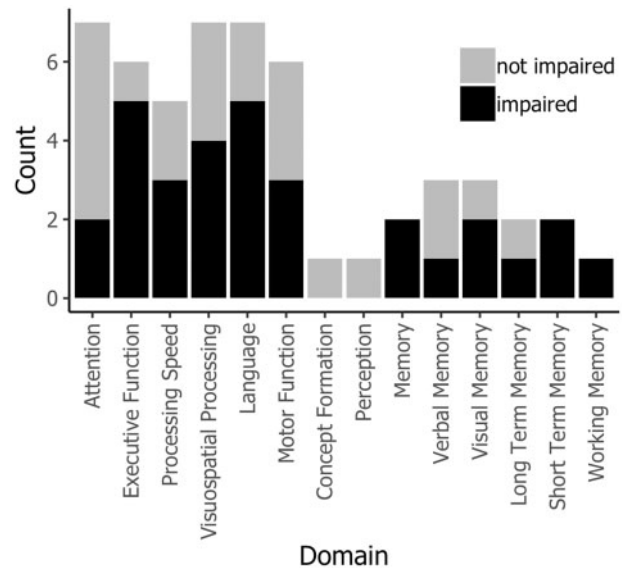


Figure 1. Profile of cognitive deficits identified in the seven meta-analyses considered in this paper. For these purposes, domains are classified as simply "impaired" or "not impaired" using conventional statistical significance. For each domain, black bars indicate the count of meta-analyses that found the domain to be impaired, and gray bars provide context in terms of the count of meta-analyses that did not find impairment. Some domains are reported more than others. For papers that report different comparison types separately (18,21,25), we show the cross-sectional data.

directly to functions. For example, Attention is a common domain name. Different measures that are typically included in the domain of Attention might assess the function of sustained attention or the function of selective attention, or they might not relate to attentional function at all, but instead measure short-term memory. Domain names vary from paper to paper. For example, what Anderson-Hanley et al. (18), Falletti et al. (20), Jim et al. (23), and Ono et al. (25) call Attention, Hodgson et al. (24) call Orientation and Attention, Jansen et al. (21) call Attention and Concentration, and Stewart et al. (22) call Simple Attention. We have therefore constructed a table of domain name synonyms (Supplementary Table 1, available online). To continue the example, we use Attention to encompass Attention, Simple Attention, Orientation and Attention, and Attention and Concentration.

Figure 1 shows the profile of impairment observed across all seven meta-analyses. For these purposes, an impairment was defined as a negative effect size (ie, performance was worse after chemotherapy than at baseline or compared with controls) for which the 95% confidence interval did not exceed zero. Most of the meta-analyses agree that there are impairments in the domains of Executive Function (five out of six find impairments), Language (five out of seven find impairments), and Memory (six out of seven find impairments for some type of memory). More disagreement exists for Processing Speed (three out of five find impairments) and Motor Function (three out of six find impairments), whereas most find no evidence of deficits in Attention (two out of five find impairments).

However, substantial disagreement exists. For every domain reported by more than two meta-analyses, at least one meta-analysis reports impairment, whereas at least one other meta-analysis reports no impairment. In Figure 2, we examine this disagreement in more detail by plotting the effect sizes reported by each meta-analysis for each domain. The only clear patterns

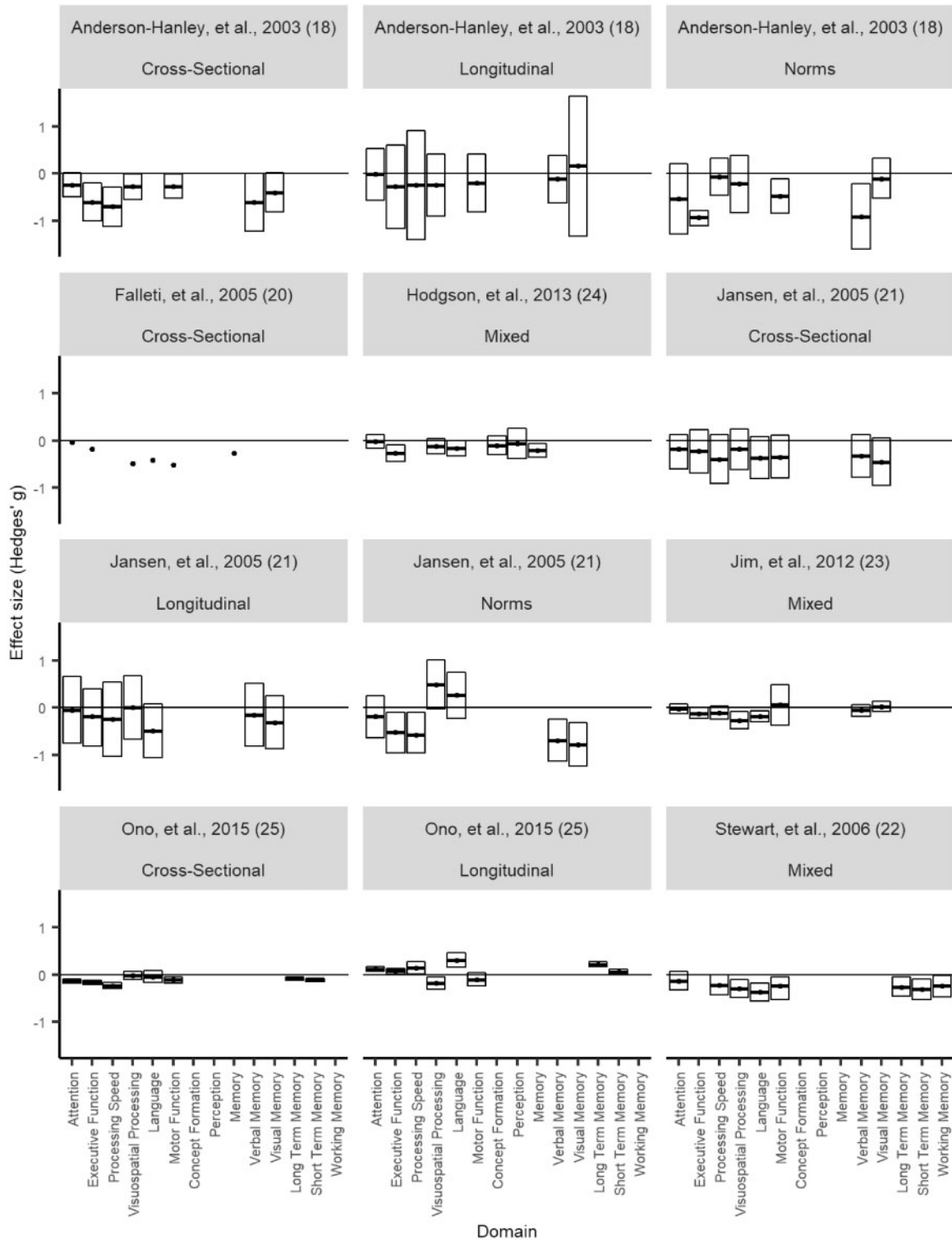


Figure 2. Effect sizes by domain for all seven meta-analyses plotted separately. Black circles denote means. Boxes denote 95% confidence intervals. Where different designs (cross-sectional, longitudinal, or comparison to population norms) are reported separately (18,21,25), these are plotted separately; data from meta-analyses that collapsed across designs are labeled “mixed.” Note that Falleti et al. (20) did not report confidence intervals.

that we can discern here are that effect sizes are generally negative (indicating impairment), and confidence intervals are generally smaller for later meta-analyses compared with earlier. As an example of discrepancies in the literature, note that Jim et al. (23) found the largest negative effect sizes for Visuospatial

Processing and Language, whereas Jansen et al. (21) found that chemotherapy patients performed better than population norms in these domains.

Each meta-analysis looks at the CRCI literature through a slightly different lens because each has different inclusion

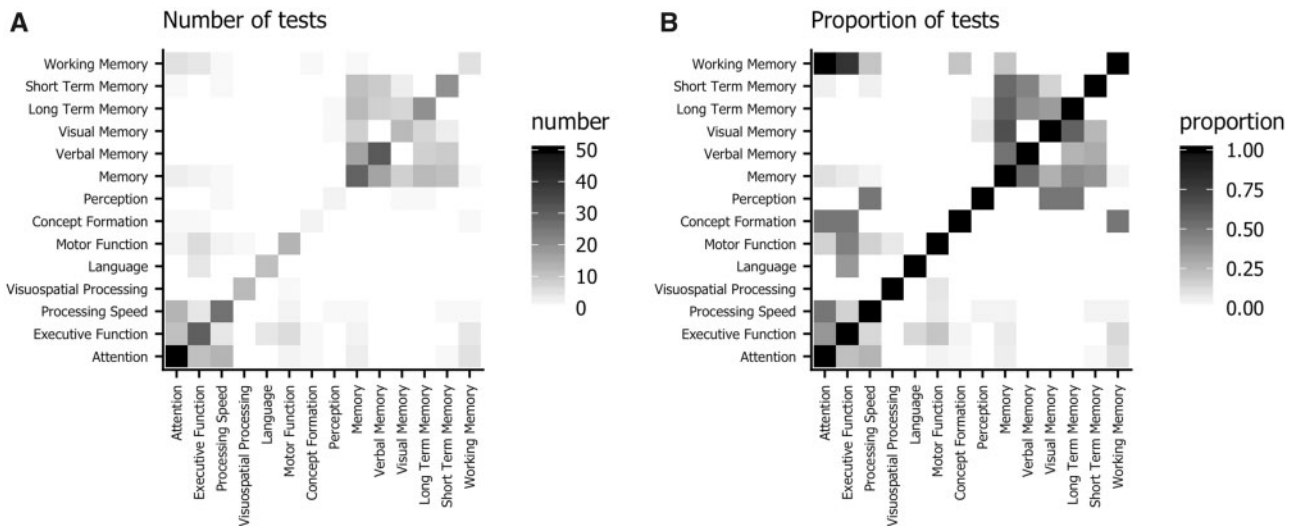


Figure 3. Confusion matrix for domains in the meta-analyses. Each row and column represents a domain. The value in each cell indicates how often a test classified into the row domain is classified into the column domain in other meta-analyses. **A)** shows the absolute number of tests, and in **B)** data are normalized by the total number of tests in the row domain. For example, the upper-left corner indicates that tests classified into Working Memory by some meta-analysis authors are often classified into Attention by other authors.

criteria and was compiled at a different time. Therefore, they include different, although overlapping, sets of studies. These differences may explain some of the disagreement. For example, Falletti et al. (20), Jim et al. (23), Ono et al. (25), and Stewart et al. (22) restricted their analyses to breast cancer patients, whereas Anderson-Hanley et al. (18), Hodgson et al. (24), and Jansen et al. (21) included studies of patients with multiple different cancer types. At least two meta-analyses (24,25) included data from patients on active treatment, whereas others included data collected more than a decade after treatment (23–25).

Different baselines may lead to different conclusions. Some of the meta-analyses break out their results separately for cross-sectional comparisons (ie, with matched control groups), longitudinal comparisons (with the patients' own pretreatment baseline), or comparisons with population norms. Figure 2 illustrates how the choice of baselines can change the results. In general, cross-sectional comparisons yield more evidence of impairment than longitudinal comparisons, suggesting that CRCI may be due in part to factors other than treatment. Some studies have shown cognitive impairments, relative to population norms, in cancer patients prior to treatment (10,28,29). If patients are already functioning below normal levels at baseline, it may be difficult to detect any additional effects of cancer therapy. The causes of this baseline impairment have not been identified. Elevated stress levels due to a cancer diagnosis are one plausible candidate (11). Biological pathways that act as risk factors for cancer may also put patients at risk for accelerated cognitive decline; examples include chronic inflammation, DNA damage or impaired DNA repair functions, hormonal alterations, and oxidative stress (30). The cross-sectional category also masks differences in the choice of matched controls, for example, healthy controls or cancer patients who underwent local surgery rather than chemotherapy.

Here we argue that an important, previously overlooked factor is that each meta-analysis has an idiosyncratic definition of the cognitive domains. First, the meta-analyses disagree as to which domains are fundamental. If all meta-analyses measured the same set of domains, the height of the stacked bars in Figure 1 would be even, but they are not. Some domains are

measured more frequently than others. All the meta-analyses report the domains of Attention, Visuospatial Processing, and Language; six of the seven report Executive Function and Motor Function; and five of the seven include Processing Speed. Hodgson et al. (24) report data for two domains that are not mentioned in the other six meta-analyses: Perception, and Concept Formation and Reasoning. Finally, the meta-analyses disagree substantially regarding the domain(s) of Memory. Three meta-analyses (18,21,23) divide Memory into Verbal and Visual Memory, whereas Ono et al. (25) and Stewart et al. (22) chose to distinguish between Long-Term and Short-Term Memory; Stewart et al. also add Working Memory to the Long-Term and Short-Term Memory domains. The remaining two meta-analyses (20,24) report only a generic Memory domain.

Even when meta-analyses agree on which domains to report, they do not necessarily agree on which neuropsychological tests measure these domains. For example, Anderson-Hanley et al. (18) classify the Paced Auditory Serial Addition Test (PASAT) as a measure of Executive Function, whereas Falletti et al. (20) classify it as a measure of Attention. Complicating matters is the fact that this disagreement occurs in the original research papers as well. Figure 3 visualizes this disagreement. Each row and column represents a domain, and the shading at the intersection indicates how often a test is classified into both domains in different meta-analyses. The left-hand panel plots the absolute number of tests, and the right-hand panel normalizes these counts by the total number of tests in the row domain. For example, there are 10 tests that are classified as Working Memory measures in some of the meta-analyses (eg, Digit Span Backwards and Arithmetic). All these tests are classified as Attention measures in other meta-analyses. The absolute number of these tests is not very large, so the upper-left corner (mirrored in the lower-right corner) of panel A) is colored light gray. However, the proportion is 1.0, so the upper-left corner of panel B) is black. The off-diagonal pattern in the upper-right corner of the plots shows the cluster of Memory tests. Tests that are classified in some meta-analyses as Short-Term Memory will show up frequently either in the Verbal Memory or generic Memory domains, and less often in the Visual Memory

domain. The cluster in the lower-left corner illustrates how the same tests are classified as Attention, Memory, or Processing Speed, depending on the meta-analysis. In addition, Working Memory tests are sometimes classified into this cluster.

Quantifying Internal Consistency in the Meta-Analyses

Given that there are multiple cognitive functions, not every neuropsychological test taps the same faculty, and these functions may be differentially affected in CRCI, it is important that tests be consistently categorized in the proper domains. To illustrate this point, assume that there are only two functions, X and Y, each of which can be measured by two tests, x_1 and x_2 , and y_1 and y_2 , respectively. Now, let's say that function X is impaired in chemotherapy patients, but function Y is not. If we conduct a meta-analysis where we use x_1 , x_2 , and y_2 to measure X, we will reduce the evidence that X is impaired, because y_2 is diluting the effect. Conversely, if we use x_1 , y_1 , and y_2 to measure Y, we might erroneously conclude that function Y is impaired. In both cases, our study may contradict another meta-analysis that categorizes the two domains properly, or a third study that used x_1 and x_2 to measure Y and y_1 and y_2 to measure X. If we think about this problem with 5–8 domains and hundreds of tests, we can see the potential for a lack of internal consistency to hinder our ability to detect real cognitive impairments that cancer survivors might be suffering from.

In this section, we attempt to quantify agreement among the meta-analyses. All the tests reported in the seven meta-analyses (18,20–25) were compiled into a matrix that counted how many times each test was classified under each domain (see [Supplementary Table 2](#), available online). We identified synonyms for tests (eg, “Booklet Category Test” and “Categories Short Booklet”) and domains (see [Supplementary Table 1](#), available online) wherever possible to minimize spurious disagreement between meta-analyses. Note that the unit of analysis here is the use of a given test in a meta-analysis. A single instance of a test in a specific original research paper may show up several times in our analysis if it was reported in multiple meta-analyses. For example, there are 80 instances of version B of the Trail Making Test (TMT-B) in the analysis, but this does not mean that 80 original research papers used the TMT-B. Our focus here is on characterizing the meta-analyses, not the underlying set of original research papers.

Although there was a consensus for the majority of tests in the database (93 of 159, 58.4%), 31 tests (19.5%) were classified into two domains, 33 tests (20.8%) were classified into three domains, and two tests (1.3%) (TMT-B and the PASAT) were classified into four domains. Although the TMT-B was primarily classified under Executive Function (67 instances), it also appeared under attention (nine instances), working memory (three instances), and memory (one instance). Although the PASAT was classified primarily as an Attention test (10 instances), it was also placed under Executive Function (two instances) as well as Working Memory (one instance) and Processing Speed (one instance). Meta-analyses agree more on the TMT-B than the PASAT. The TMT-B is classified as an Executive Function test 83.8% of the time, whereas the PASAT is classified as Attention 71.4% of the time. However, the greater sample size for the TMT-B (80 vs 14) makes the comparison with the PASAT difficult. The test generating the most disagreement might be the Digit Symbol test. It shows up in only three domains, whereas the TMT-B and PASAT show up in four.

However, as noted, the TMT-B is mostly used to measure Executive Function, and only rarely classified into the other three domains; the same goes for the PASAT and Attention. The Digit Symbol test, in contrast, is classified more evenly across domains (43.1% Processing Speed, 32.3% Attention, and 24.6% Executive Function).

To account for the number of times a test shows up in the database, the number of different domains it is classified into, and the degree to which classifications cluster into one domain or are more evenly distributed, we employed the χ^2 statistic. The distribution of counts for each test can then be summarized with a one-way χ^2 statistic, indicating the degree of unevenness in the count distribution. A test always classified under the same domain would have a high χ^2 value, whereas a test that was randomly classified would have a low χ^2 value. However, given the same underlying distribution, a test that is reported more often (eg, TMT-B) will also have a higher χ^2 value than one that is reported rarely (eg, the PASAT). Therefore, for each test, we also computed its ideal χ^2 statistic: the value that would have been generated if the test were always classified into a single domain. This gives us two distributions of χ^2 statistics: the ideal and the actual. The less disagreement there is among meta-analyses, the closer the actual distribution will be to the ideal.

The top row of [Figure 4](#) plots the histograms for the two distributions. The actual distribution (in black) is noticeably more peaked and shifted to the left (low values) relative to the ideal distribution (blue). This illustrates that there is often substantial disagreement between meta-analysis authors on how to classify individual tests, which at least partially explains the variability illustrated in [Figures 1](#) and [2](#). This disagreement is not just a product of the meta-analyses. If we look at how the authors of the original research papers classify tests, we see a similar lack of agreement, as shown in the bottom row of [Figure 4](#).

Discussion

As we noted earlier, meta-analysis should help bring clarity to a scientific field by replacing somewhat subjective qualitative summaries with quantitative rigor. As [Figures 1](#) and [2](#) illustrate, however, meta-analyses of CRCI do not show consistent findings. What domains are impaired in CRCI? Motor Control? Processing Speed? Attention? Short-Term Memory? Long-Term Memory? The answer varies depending on which meta-analysis is consulted. Although this may be due in part to different study selection criteria, we suggest that disagreement among authors on how the various domains should be measured (as well as what the basic cognitive domains are) is a major barrier to understanding CRCI. This disagreement is quantified in [Figure 4](#). The lack of internal consistency in the field makes it difficult to compare across studies and reduces the utility of meta-analysis.

One solution would be to establish common guidelines on how to classify neuropsychological tests for the purposes of meta-analysis. For example, the International Cognition and Cancer Task Force released recommendations for common neuropsychological measures, study design, and analysis procedures for primary studies (12). As part of a broader movement toward improved statistical sophistication and data reporting requirements for original studies (31) and best practices for conducting and reporting meta-analyses (32,33), a task force could

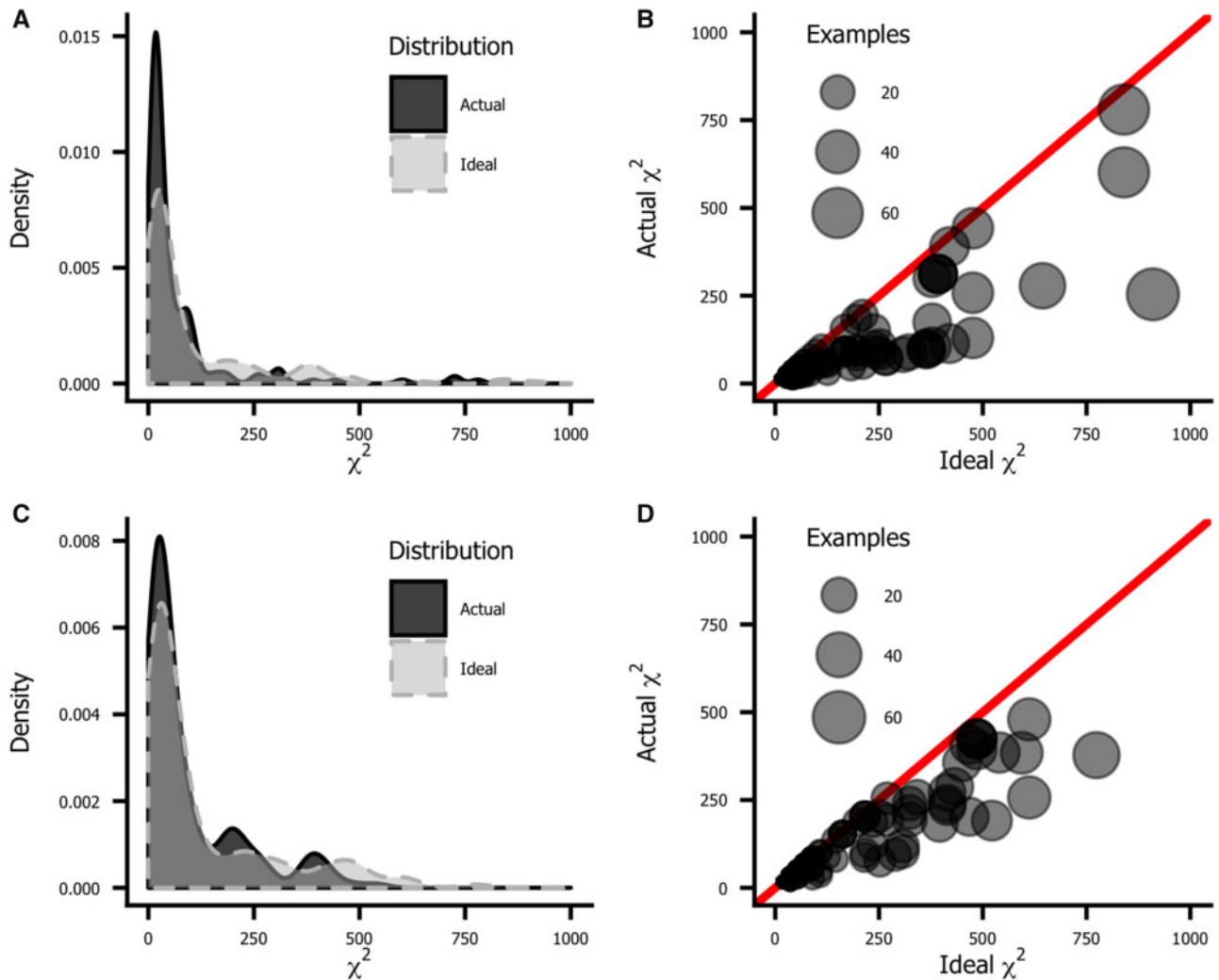


Figure 4. Agreement among authors on how to assign tests to domains, as measured by the χ^2 statistic. The ideal χ^2 statistic is the value that would be observed if there were complete consensus on every test. The less consensus there is, the lower the χ^2 statistic. The ideal χ^2 varies from test to test depending on the number of observations; more frequently used tests naturally produce greater χ^2 values than less frequently used tests. **A)** and **B)** plot agreement among meta-analyses, and **C)** and **D)** plot agreement among original research papers. **A)** and **C)** show histograms of the ideal (light gray) and actual (dark gray) χ^2 statistics. The actual distributions are peakier and left-shifted relative to the ideal, indicating a lack of agreement among authors. **B)** and **D)** are scatterplots of actual χ^2 values plotted against the ideal. Symbol size reflects the number of instances of a given test in the database. If there were complete consensus, all symbols would fall on the diagonal line of equality. The further a symbol lies below this line, the more disagreement there is about which domain to place it in. More frequently used tests tend to be inconsistently classified.

be convened to recommend standard guidelines for assigning tests to domains.

We suggest that these guidelines should be informed by the way that cognitive domains are conceptualized in the experimental fields of cognitive psychology and neuroscience (34). Consider span tests, such as Digit Span Forwards, widely used in neuropsychological batteries such as the Wechsler Adult Intelligence Scale and Wechsler Memory Scale series. In our dataset, Digit Span Forwards was the second most frequently used Attention test. Furthermore, researchers universally agreed that it should be classified under Attention. Yet from a theoretical point of view, Digit Span Forwards has nothing to do with attention beyond the minimal sustained attention necessary to sit through the task. A number of authors in the neuropsychological literature have made the point that working memory tasks such as Digit Span Forwards are neither theoretically nor empirically related to attention (35–37); even the

standard four-factor model of the WAIS IV places Digit Span in a working memory factor (38). Improperly classifying Digit Span Forwards and other working memory tests under Attention is another barrier to gaining a concrete understanding of the challenges faced by CRCI patients.

To understand, predict, and treat CRCI, we need to be able to measure the profile of cognitive impairments that make up this syndrome. From the substantial research literature on CRCI, we should be able to state clearly which cognitive domains are most affected and which are spared. Unfortunately, there is still substantial disagreement among researchers. In this paper, we have demonstrated that this disagreement stems in part from a lack of internal consistency. Neither empirical researchers nor the authors of meta-analyses agree on which cognitive domains should be measured, nor which tests measure which domains. Although CRCI research faces many important scientific challenges without clear solutions, this is one problem that can be

solved by better communication among scientists. This issue is not particular to CRCI; both internal and external consistency are general challenges for neuropsychological research (37). A set of agreed on standards for assigning existing neuropsychological tests to domains would improve the ability to interpret existing findings, facilitate comparison across meta-analyses, and accelerate the drive to understand and ameliorate CRCI. Furthermore, such a process could become a model for improving the use and interpretation of neuropsychological tests in general, much as the CNTRICS program (39,40) serves as a model for improving the development and selection of tests for specific disorders.

Notes

Affiliations of authors: National Cancer Institute, Rockville, MD (TSH, MT, IMG, KAD).

The authors would like to thank Sashant Palli for assistance with data entry. We would also like to thank Catherine Alfano, Paige Green, Lynne Padgett, and Jerry Suls for invaluable advice. No outside funding was involved in this project. The authors have no conflicts of interest directly related to this work to disclose.

References

- Miller KD, Siegel RL, Lin CC, et al. Cancer treatment and survivorship statistics, 2016. *CA Cancer J Clin*. 2016;66(4):271–289.
- Rodin G, Ahles TA. Accumulating evidence for the effect of chemotherapy on cognition. *J Clin Oncol*. 2012;30(29):3568–3569.
- Schagen SB, van Dam FS, Muller MJ, Boogerd W, Lindeboom J, Bruning PF. Cognitive deficits after postoperative adjuvant chemotherapy for breast carcinoma. *Cancer*. 1999;85(3):640–650.
- Donovan KA, Small BJ, Andrykowski MA, Schmitt FA, Munster P, Jacobsen PB. Cognitive functioning after adjuvant chemotherapy and/or radiotherapy for early-stage breast carcinoma. *Cancer*. 2005;104(11):2499–2507.
- Jenkins V, Shilling V, Deutsch G, et al. A 3-year prospective study of the effects of adjuvant treatments on cognition in women with early stage breast cancer. *Br J Cancer*. 2006;94(6):828–834.
- Castellon SA, Ganz PA, Bower JE, Petersen L, Abraham L, Greendale GA. Neurocognitive performance in breast cancer survivors exposed to adjuvant chemotherapy and tamoxifen. *J Clin Exp Neuropsychol*. 2004;26(7):955–969.
- Srisurapanont M, Suttajit S, Eurviriyanyukul K, Varnado P. Discrepancy between objective and subjective cognition in adults with major depressive disorder. *Sci Rep*. 2017;7(1):3901.
- Yoon BY, Lee J-H, Shin SY. Discrepancy between subjective and objective measures of cognitive impairment in patients with rheumatoid arthritis. *Rheumatol Int*. 2017;37(10):1635–1641.
- Shilling V, Jenkins V, Trapala IS. The (mis)classification of chemo-fog – methodological inconsistencies in the investigation of cognitive impairment after chemotherapy. *Breast Cancer Res Treat*. 2006;95(2):125–129.
- Cimprich B, Reuter-Lorenz P, Nelson J, et al. Prechemotherapy alterations in brain function in women with breast cancer. *J Clin Exp Neuropsychol*. 2010;32(3):324–331.
- Berman MG, Askren MK, Jung M, et al. Pretreatment worry and neurocognitive responses in women with breast cancer. *Health Psychol*. 2014;33(3):222–231.
- Wefel JS, Vardy J, Ahles T, Schagen SB. International Cognition and Cancer Task Force recommendations to harmonise studies of cognitive function in patients with cancer. *Lancet Oncol*. 2011;12(7):703–708.
- Nelson WL, Suls J. New approaches to understand cognitive changes associated with chemotherapy for non-central nervous system tumors. *J Pain Symptom Manage*. 2013;46(5):707–721.
- Land VJ, Sutow WW, Fernbach DJ, Lane DM, Williams TE. Toxicity of L-asparaginase in children with advanced leukemia. *Cancer*. 1972;30(2):339–347.
- Peterson LG, Popkin MK. Neuropsychiatric effects of chemotherapeutic agents for cancer. *Psychosomatics*. 1980;21(2):141–153.
- Ahles TA, Root JC, Ryan EL. Cancer- and cancer treatment-associated cognitive change: an update on the state of the science. *J Clin Oncol*. 2012;30(30):3675–3686.
- Ahles TA, Saykin A. Cognitive effects of standard-dose chemotherapy in patients with cancer. *Cancer Invest*. 2001;19(8):812–820.
- Anderson-Hanley C, Sherman ML, Riggs R, Agocha VB, Compas BE. Neuropsychological effects of treatments for adults with cancer: a meta-analysis and review of the literature. *J Int Neuropsychol Soc*. 2003;9(7):967–982.
- Jansen CE, Miaskowski CA, Dodd MJ, Dowling GA. A meta-analysis of the sensitivity of various neuropsychological tests used to detect chemotherapy-induced cognitive impairment in patients with breast cancer. *Oncol Nurs Forum*. 2007;34(5):997–1005.
- Falletti MG, Sanfilippo A, Maruff P, Weih L, Phillips K-A. The nature and severity of cognitive impairment associated with adjuvant chemotherapy in women with breast cancer: a meta-analysis of the current literature. *Brain Cogn*. 2005;59(1):60–70.
- Jansen CE, Miaskowski C, Dodd M, Dowling G, Kramer J. A metaanalysis of studies of the effects of cancer chemotherapy on various domains of cognitive function. *Cancer*. 2005;104(10):2222–2233.
- Stewart A, Bielajew C, Collins B, Parkinson M, Tomiak E. A meta-analysis of the neuropsychological effects of adjuvant chemotherapy treatment in women treated for breast cancer. *Clin Neuropsychol*. 2006;20(1):76–89.
- Jim HSL, Phillips KM, Chait S, et al. Meta-analysis of cognitive functioning in breast cancer survivors previously treated with standard-dose chemotherapy. *J Clin Oncol*. 2012;30(29):3578–3587.
- Hodgson KD, Hutchinson AD, Wilson CJ, Nettelbeck T. A meta-analysis of the effects of chemotherapy on cognition in patients with cancer. *Cancer Treat Rev*. 2013;39(3):297–304.
- 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):1–19.
- 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.
- Lindner OC, Phillips B, McCabe MG, et al. A meta-analysis of cognitive impairment following adult cancer chemotherapy. *Neuropsychology*. 2014;28(5):726–740.
- Wefel JS, Lenzi R, Theriault RL, Davis RN, Meyers CA. The cognitive sequelae of standard-dose adjuvant chemotherapy in women with breast carcinoma. *Cancer*. 2004;100(11):2292–2299.
- Hermelink K, Untch M, Lux MP, et al. Cognitive function during neoadjuvant chemotherapy for breast cancer. *Cancer*. 2007;109(9):1905–1913.
- Mandelblatt JS, Hurria A, McDonald BC, et al. Cognitive effects of cancer and its treatments at the intersection of aging: what do we know; what do we need to know? *Semin Oncol*. 2013;40(6):709–725.
- Cumming G. The new statistics why and how. *Psychol Sci*. 2014;25(1):7–29.
- Lakens D, Hilgard J, Staaks J. On the reproducibility of meta-analyses: six practical recommendations. *BMC Psychol*. 2016;4(1):24.
- Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ*. 2010;340(Feb05):c221.
- Horowitz TS, Suls J, Treviño M. A call for a neuroscience approach to cancer-related cognitive impairment. *Trends Neurosci*. 2018;41(8):493–496.
- Mirsky AF, Anthony BJ, Duncan CC, Ahearn MB, Kellam SG. Analysis of the elements of attention: a neuropsychological approach. *Neuropsychol Rev*. 1991;2(2):109–145.
- Schmidt M, Trueblood W, Merwin M, Durham RL. How much do ‘attention’ tests tell us? *Arch Clin Neuropsychol*. 1994;9(5):383–394.
- Johnstone B, Erdal K, Stadler MA. The relationship between the Wechsler Memory Scale—revised (WMS-R) attention index and putative measures of attention. *J Clin Psychol Med Settings*. 1995;2(2):195–204.
- Reynolds MR, Ingram PB, Seeley JS, Newby KD. Investigating the structure and invariance of the Wechsler Adult Intelligence Scales, Fourth Edition, in a sample of adults with intellectual disabilities. *Res Dev Disabil*. 2013;34(10):3235–3245.
- Carter CS, Barch DM. Cognitive neuroscience-based approaches to measuring and improving treatment effects on cognition in schizophrenia: the CNTRICS initiative. *Schizophr Bull*. 2007;33(5):1131–1137.
- Moore H, Geyer MA, Carter CS, Barch DM. Harnessing cognitive neuroscience to develop new treatments for improving cognition in schizophrenia: CNTRICS selected cognitive paradigms for animal models. *Neurosci Biobehav Rev*. 2013;37(9):2087–2091.