Neuropsychological effects of treatments for adults with cancer: A meta-analysis and review of the literature

CAY ANDERSON-HANLEY,[1] MARNE L. SHERMAN,[2] RAINE RIGGS,[3] V. BEDE AGOCHA,[4]
AND BRUCE E. COMPAS[5]

[1]Skidmore College and the Cancer Center at Glens Falls Hospital, Glens Falls, New York
[2]University of Missouri–Kansas City and The Cancer Institute, Kansas City, Missouri
[3]Brown University School of Medicine, Providence, Rhode Island
[4]University of Connecticut, Farmington, Connecticut
[5]Vanderbilt University, Nashville, Tennessee

Abstract

A meta-analysis was conducted to evaluate possible neuropsychological effects of treatments for cancer in adults. A search revealed 30 studies, encompassing 29 eligible samples, and leading to inclusion of a total of 838 patients and control participants. A total of 173 effect sizes (Cohen's *d*) were extracted across 7 cognitive domains and as assessed in the literature via 3 methods of comparison (post-treatment compared with normative data, controls, or baseline performance). Statistically significant negative effect sizes were found consistently across both normative and control methods of comparison for executive function, verbal memory, and motor function. The largest effects were for executive function and verbal memory normative comparisons (−.93 and −.91, respectively). When limiting the sample of studies in the analyses to only those with relatively "less severe" diagnoses and treatments, the effects remained. While these results point toward some specific cognitive effects of systemic cancer therapies in general, no clear clinical implications can yet be drawn from these results. More research is needed to clarify which treatments may produce cognitive decrements, the size of those effects, and their duration, while ruling out a wide variety of possible mediating or moderating variables. (*JINS*, 2003, *9*, 967–982.)

Keywords: Meta-analysis, Cancer, Chemotherapy, Executive function, Verbal memory, Motor function

INTRODUCTION

Recent estimates indicate that there are approximately 8.9 million cancer survivors in the United States, representing nearly 3% of the population (NCI, 1998). Virtually all survivors have been treated with either surgery, radiation, chemotherapy, biologics, hormonal therapies, or some combination of these. As treatments for cancer have improved in efficacy and length of survival has increased for many patients, extensive research has been done to document the acute and long-term side effects of these treatments (e.g., Bonadonna et al., 1985; Brundage, 1997; Fisher et al., 1994; Petros, 2002; Saphner et al., 1991; Videtic, 2001).

While researchers have focused their attention on physical side effects, another potentially important side effect, impaired neuropsychological functioning, has received less attention. Possible neuropsychological side effects of these treatments may include difficulty concentrating, impaired verbal and visual memory, difficulty organizing information, decreased motor skills, and language problems (e.g., word finding difficulty). Health care professionals and researchers have become increasingly aware of possible cognitive effects through the anecdotal reports of some cancer survivors. Some patients have reported that problems in cognitive functioning have precluded a smooth transition to life after cancer (e.g., problems resuming complex tasks at work or multi-tasking at home). These neuropsychological impairments are so well known among cancer survivors that they have been described in several cancer patient newsletters and have been dubbed "chemobrain" (Mann, 1999[1]). Thus, clinicians and researchers concerned with quality of life issues relevant to patients have begun to turn their attention to this matter.

Reprint requests to: C. Anderson-Hanley, Department of Psychology, Skidmore College, Saratoga Springs, NY 12866. E-mail: chanley@skidmore.edu

[1]See www.mynewspirit.com/chemobrain.htm.

Scientific studies of the chemobrain phenomenon began in the 1970s (e.g., Weiss et al., 1974a, 1974b), and have appeared with increasing frequency in the literature since then. A number of reviews have been published regarding the possible neuropsychological effects of a variety of cancer treatments, but interpretations of the findings have not been definitive (e.g., Ahles & Saykin, 2001; Bender et al., 2001; Fleishman & Kalash, 1998; Ganz, 1998; Meyers, 2000; Meyers & Scheibel, 1990; Meyers et al., 1994; Olin, 2001; Peterson & Popkin, 1980; Redd et al., 1991; Silberfarb, 1983; Silberfarb & Oxman, 1988; Tope et al., 1993; Trask et al., 2000; Troy et al., 2000; Walch et al., 1998). Earlier reviews had fewer studies to work with and lacked the methodological rigor that is becoming more common in current research on possible neuropsychological effects. Ganz (1998) notes that there are some equivocal results and points to the possibility of a dose effect. Redd et al. (1991) interpreted the existing literature as suggesting cognitive deficits as a result of treatment, but also expressed concern about the tools used in the measurement of these effects. Some recent reviews have focused on effects within a specific diagnosis, such as Olin's review of breast cancer (2001), while other reviews have focused on a specific treatment, such as the review by Trask and colleagues of Interferon (2000), or the review by Troy and colleagues of Cisplatin (2000). The most recently published qualitative review of the literature, by Ahles and Saykin (2001), concluded that standard-dose chemotherapy is associated with neuropsychological impairments in a subgroup of adult cancer survivors. The authors concluded that the impairments include subtle decrements in memory and concentration and that these subtle changes can have a significant impact on cancer survivors' quality of life.

The present review is different because it is a quantitative review of all available studies. None of the above reviews was quantitative in nature and therefore none could address the question of the *size* of an effect cancer treatments might be having. Additionally, a meta-analysis has strength in its numbers, in that patients from across many studies are pooled, allowing for an increase in power to detect an effect when one exists.

The present meta-analysis attempts to clarify the variety of conclusions from qualitative reviews by bringing together the results of quantifiable inquiries into clearly defined cognitive domains measured with standardized assessment tools. This meta-analysis examined neuropsychological effects of cancer treatments across different types of cancers and given different types of treatments. While cancer treatments are not equal and therefore unlikely producing a unitary effect on cognitive functioning, the limited availability of research in this area leads to this synthesis in order to provide a preliminary and general sense of possible effects. It is anticipated that additional research will need to be done so that more definitive conclusions can be drawn about the specific cognitive domains that each type of therapy affects and the potential mechanism(s) that create any cognitive deficits.

If any effects do exist in this first analysis of existing literature, we wanted to know for which domain(s) of cognitive functioning there exists an effect, and also how big that effect is. All effects were converted to a common metric (Cohen's *d*), so that outcomes could be compared quantitatively across measures and domains of neuropsychological function. Results of studies were pooled according to the three types of research designs that have been employed in the existing literature. Two of the designs used were between subjects designs, one allowing for comparisons of post-treatment results with control subjects and the other allowing for comparisons with normative data. The third design used was a within subjects design, allowing for comparisons of post-treatment results with a subjects' own baseline performance. Thus, this latter design is longitudinal and the former two designs are cross-sectional.

Synthesizing these results is an important advantage of meta-analysis since these three methods of comparison might be expected to produce somewhat different results. Comparisons of post-treatment results with baseline might be considered the ideal method since using a subject's own baseline controls many extraneous variables (e.g., holds constant prior experience and education). However, significant practice and learning effects are common with many tests (Dikmen et al., 1999; McCaffrey et al., 2000; Salinsky et al., 2001; Temkin et al., 1999). For example, it can be problematic if a testee is asked to recall the same list of words on a memory task that had been used during a previous test session, or if a testee is asked to repeat a particular procedure with which they have already been familiarized, such as connecting dots in order. A relatively negligible change from pre- to post-test might not appear to be a real decline, but it could be a diminution of function counterbalanced and obscured by practice effects (which normally lead to a substantial increase from pre- to post-test; Dikmen et al., 1999; McCaffrey et al., 2000; Salinsky et al., 2001; Temkin et al., 1999). The design that uses comparisons with control subject performance also has its benefits (e.g., eliminates practice and learning effects, while providing a theoretically comparable contrast group such as persons of similar ages and education from the same geographic regions). Patients with a similar diagnosis who are not prescribed a particular treatment could be said to serve as excellent control subjects; however, they might differ on some third variable (e.g., no treatment controls might differ on overall health status, such that their performance status may have been too low to tolerate chemotherapy and so it was withheld). Comparisons with healthy controls and normative data are also helpful, but a drawback is that patients may differ on some important third variable (e.g., fatigue, distress level). We might expect the greatest effects to show up in normative comparisons since normative samples typically contain individuals with a wider range of education and experience. One strength of the present meta-analysis is that we will be able to contrast all three methods of comparison and look for convergence. Ultimately, the meta-analysis allows for

quantification of the magnitude of any effects that are significantly different from zero.

For this meta-analysis, we selected all identifiable research on the neuropsychological effects of systemic therapies for adults with cancer. While there are numerous published studies on the neuropsychological sequelae of treatments for childhood cancer, we chose not to include those studies in this analysis. The effects of cancer treatments on the developing brain are thought to be significantly different from the effects on the adult brain (for reviews see Butler, 2002; Copeland et al., 1985; Eiser, 1998; Gotay, 1987). We also chose to exclude studies that focused exclusively on the effects of treatments for brain cancers. Treatments for these cancers often involve brain irradiation and surgical interventions, in addition to direct effects on brain tissue secondary to the lesions (Roman & Sperduto, 1995; Weitzner & Meyers, 1997). In such cases, changes in neuropsychological functioning would be expected and could have skewed the results of this study. Similarly, we excluded studies in which all patients received some direct treatment to the brain (e.g., cranial irradiation). While some studies in our sample include a percentage of patients with gliomas, central nervous system (CNS) lesions or brain irradiation, these studies were retained since the samples also included other cancer diagnoses and treatments.

In addition to examining the available literature for an overall pattern of results, we also considered the role of possible mediating and moderating variables. We hypothesized that the severity of diagnosis and the intensity of treatment would impact any effect found on neuropsychological functioning. Thus, we conducted analyses that included only those studies in which diagnoses were relatively less severe (e.g., non-metastatic) and treatments were relatively less intense (e.g., excluding any brain irradiation or bone marrow transplant, BMT). We conducted these additional sub-group analyses to evaluate whether these two groups of studies produced similar results and thus, whether or not the "more severe" studies might be having an untoward effect on overall effect sizes.

METHODS

Literature Search and Criteria for Inclusion

We performed a literature search of the PsychInfo, Medline and CancerLit databases. Key words included: *cancer*, *neoplasm*, *oncology*, *oncologic*, *chemotherapy*, *systemic treatment*, *drug effects*, *behavior*, *cognition*, *cognitive*, *tests*, *measures*, *assessment*, *neuropsychology*, and *neuropsychological*. Pre-print, unpublished, and file-drawer papers were requested via the Division 38 (Health Psychology) of the American Psychological Association list-serve and via email to authors who had already published on this topic. Research articles retrieved in this manner were also inspected for relevant references to locate additional articles.

Nearly 100 articles were located that addressed the issue of cancer treatment and neuropsychological functioning. To be included in analyses, the report needed to contain original study data (e.g., not a review), the sample needed to be adult, diagnoses could not be entirely metastatic or gliomas, and treatment could not involve brain irradiation of the whole sample. To qualify for inclusion, it was necessary that the study had reported quantitative measurement or an inferential statistic regarding some aspect of neuropsychological functioning. If not, an attempt was made to retrieve such data (e.g., means, standard deviations) directly from the study's author(s). Thirty of the studies that had been identified met criteria and were included in this meta-analysis (leading to 29 samples: two samples came from Ahles et al., 2002, and in two instances a single sample spanned two articles by the same first author: Kaasa et al., 1988a, 1988b; and Walker et al., 1996, 1997). Excluded studies did not contain quantitative data or statistics that could be analyzed. Table 1 displays information on the studies that were included in our analyses.

Recruitment of Participants into Studies

It is important to consider how participants came to be enrolled in the studies included in this meta-analysis, since a selection bias in favor of clinically identified participants could artificially inflate any results. A review of the methods used by each study included here revealed that, in general, the researchers used exhaustive methods to identify and recruit all eligible participants at their sites (based on tumor and/or treatment type, not based on referral by self or other due to cognitive concerns). One study did not clarify the method of selection, but also did not specify a referral basis, and one study specified that only one participant was referred because of concerns about mental status changes.

Cognitive Domains and Measures

A variety of cognitive domains were assessed across studies and many different measures were utilized to tap these domains. The following are the domains that could be compiled based on measures used in the literature: *Attention* was measured by six tests: (1), Trails A; (2) Wechsler Adult Intelligence Scale (WAIS)–Digit Span; (3) High Sensitivity Cognitive Screen (HSCS)–Attention; (4) Dementia Rating Scale (DRS)–Attention; (5) Automated Performance Test System (APTS)–Sternberg Test and Simple Reaction Time; (6) Stroop–A. *Information processing speed* was measured by 5 tests: (1) WASI–Digit Symbol; (2) Monroe-Sherman Reading Comprehension; (3) Zazzo's Attention Test–Speed; (4) APTS–Code Substitution; (5) Computer Drug Research System (CDRS)–Number Vigilance and Memory Scanning. *Verbal memory* was measured by 8 tests: (1) Wechsler Memory Scale (WMS)–Logical Memory; (2) HSCS–Memory; (3) DRS–Memory; (4) Rey Auditory Verbal Learning Test (RAVLT); (5) Memory Scanning Test; (6) Concept Shifting Test (CST)–A and B; (7) California Verbal Learning Test (CVLT); (8) Buschke Selective Re-

Table 1. Empirical studies included that examined cognitive effects of systemic treatments

First author	Year	Dx	Tx	n	Age	Weeks from dx/tx
Ahles	2002	Breast/Ly	chemo	70	56	260
Peace	2002	Breast & Bowel	chemo vs. sx	55	58	64
Capuron*	2001	Renal	IFN-α	7	57	4
Riggs	2001	Breast	chemo	13	47	52
Brezden	2000	Breast	chemo	40	46	110
Mayerhofer	2000	Ovarian	chemo	28	—	0
Schagen	1999	Breast	chemo	39	47	109
Van Dam	1998	Breast	chemo/RT	34	45	110
Walker*	96/97	Colorectal	rIL/chemo	9	56	12
Ahles*	1996	Breast/Hem	BMT	34	39	3
Van Oosterhout*	1996	Lung	chemo/TBI	49	65	300
Komaki*	1995	Lung	chemo/PCI	21	61	56
Meyers*	1995	Lung	chemo/RT/PCI	25	55	–
Pavol*	1995	CML	IFN-α	16	49	132
Wieneke	1995	Breast	chemo	28	42	24
Ahles*	1994	Lung	chemo/TBI	17	58	3
Archibald*	1994	Glioma	chemo/RT	16	35	78
Cull*	1994	Lung	chemo/PCI	64	61	110
Meyers*	1994	Mix (Leuk, Ly, My)	BMT	21	38	32
Andrykowski*	1992	Hem/Ly/My	BMT	55	36	100
Caraceni*	1992	Met Melan (non-CNS)	IL-2	7	42	1
Meyers*	1992	Met Colorectal	chemo/bio	19	54	141
Parth*	1989	—	BMT/TBI	20	26	33
Tucker*	1989	Hem & Ly	chemo/TBI	22	32	572
Kaasa	1988	Lung	chemo/RT	31	62	5
Denicoff*	1987	Met mix	IL-2	38	50	1
Oxman*	1980	Met mix	chemo	10	53	4
Silberfarb*	1980	Mix/CNS	chemo/RT	50	59	0

* = classified as high severity (diagnoses and/or treatments) for subgroup analyses. Chemo = chemotherapy; TBI = total brain irradiation; BMT = bone marrow transplantation; Hem = hematologic cancers; Ly = lymphoma; My = myeloma; Leuk = leukemia; RT = radiation therapy; IL-2 = interleukin-2; IFN-α = interferon-alpha; met = metastatic; Melan = melanoma; PCI = prophylactic cranial irradiation; CML = chronic myelogenous leukemia; sx = surgery; CNS = central nervous system; bio = biologics; rIL = recombinant interleukin.

minding Task. *Visuospatial memory* was measured by three tests: (1) Benton Visual Retention Task; (2) Rey-Osterrieth Complex Figure–Recall; (3) WMS–Non-Verbal Memory. *Visuospatial skill* were measured by 5 tests: (1) Rey-Osterrieth Complex Figure–Copy; (2) HSCS–Spatial; (3) Road Map Sense Test; (4) APTS–Pattern Comparison and Manikin Test; (5) WAIS–Block Design. *Executive function* was measured by nine tests: (1) Controlled Oral Word Association; (2) Trails B; (3) Design Fluency; (4) HSCS–Planning; (5) Wisconsin Card Sort Test (WCST); (6) Stroop–C/Interference; (7) MST–C; (8) Paced Auditory Serial Addition Test (PASAT); (9) Similarities. *Psychomotor skill* was measured by four tests: (1) Grooved Pegboard; (2) Finger Tapping; (3) DRS–Construction; (4) APTS–Tapping.

Methods of Comparison

Three research designs have been used in evaluating possible neuropsychological effects of cancer treatments: (1) treatment group outcomes compared with published *normative* data; (2) treatment group outcomes compared with *control* group outcomes; and (3) treatment group outcomes compared with subjects' own *baseline* (pre-treatment) scores. In the second type of contrast, control groups were comprised of either healthy age-matched controls or individuals with a similar diagnosis, but who were receiving some presumably more benign treatment (e.g., radiation only). These two subgroups were subsequently compared to rule out possible differences.

Statistical Analysis

Findings from each study were assigned to a cognitive domain category and converted to Cohen's d using a random effects model (Shadish & Haddock, 1994). Cohen's d is interpreted as the number of standard deviations the average cancer treatment group member differs from the average individual in the control or normative sample, or from her/his own baseline score in each of the cognitive domains. When means and standard deviations were not available, effect sizes were calculated from other reported statistics (e.g., p values, percentages) using the methods

described by Rosenthal and DiMatteo (2001), and Cooper and Hedges (1994).

Effect sizes were assigned such that negative values indicated poorer performance in the treatment group (relative to the control group, relative to the normative sample, or relative to subjects' prior performance, as appropriate). Allowing for a conservative bias and in order to protect against Type I errors, we assigned a value of zero to studies that reported nonsignificant results, but which did not include sufficient data to specifically estimate effect sizes. Cohen's (1988) guidelines for interpreting effect sizes suggest that absolute values less than .2 indicate a negligible effect, values between .2 and .5 indicate a small effect, values between .5 and .8 indicate a medium effect, and values of .8 and greater indicate a large effect.

If more than one measurement was made of a particular cognitive domain in a single study, an average of the effect sizes was used in the final analyses. This averaging of multiple within-study effect sizes yields one effect size per cognitive domain (per method) for each study and limits the degree of statistical nonindependence of the results. Some overlap in the reporting of results was inevitable since many studies measured more than one cognitive domain and used multiple methods of comparison. Given the variability in sample sizes among the studies included, effect sizes were weighted by the sample size. The combining, weighting, confidence interval, and statistical significance calculations for groups of effect sizes were computed via the Comprehensive Meta-Analysis computer program (Version 1.25; Borenstein & Rothstein, 2002).

A significance level of .01 is inferred when the 99% confidence limit does not cross zero (Shadish & Haddock, 1994). Additionally, for each mean effect size, a "fail-safe *N*" was calculated (Rosenthal, 1991). This value estimates the number of unpublished, nonsignificant studies of comparable or larger sample sizes that would have to exist for the obtained probability value to be rendered nonsignificant. Cut-offs are computed for each effect based on the number of studies used in calculating the effect (see below).

RESULTS

Included in analyses were data from 838 patients from 29 samples (Table 1). The mean age of the patients was 49 years ($SD = 10.3$) and the average amount of time from either diagnosis or treatment was 86 weeks ($SD = 124.1$). See Table 2 for pooled means and standard deviations of some of the most commonly reported measures.

From the 30 reports with sufficient data, 173 effect sizes (Cohen's *d*) were extracted across seven cognitive domains (attention, information processing, verbal memory, visuospatial memory, visuospatial skill, executive function, and psychomotor skill; Table 3). As can be seen in Table 3, these effect sizes were categorized by one of the three methods used to evaluate the patients' neuropsychological functioning in the study or *post hoc* by the first author (e.g., comparing patient's outcomes with a published normative

Table 2. Pooled means and standard deviations for the most commonly reported measures

Test	Pooled mean	Pooled SD	n studies	n total
Digit Span	11.96	2.17	5	92
Trails A	31.27	14.67	6	202
Dig Symbol (scaled score)	10.23	2.73	4	136
Dig Symbol (raw score)	55.40	9.84	6	162
CVLT–immed	10.25	2.90	2	71
CVLT–delayed	10.65	2.95	2	71
Rey-O copy	34.74	1.70	3	86
Rey-O delay	27.61	7.60	4	102
Trails B	108.29	49.50	12	378
COWA/FAS	41.90	12.11	4	118
Peg Board–dominant	61.97	8.52	2	52

Note. In several cases, more studies utilized these tests, but only the above studies reported data that could be reported as means and standard deviations.

group's performance, a control group's performance, or their own baseline prior to treatment).

Effect sizes were weighted and combined (Shadish & Haddock, 1994) to yield a single effect size for each of the seven cognitive domains across each of the three methods used to evaluate patients' neuropsychological functioning post-treatment (norm comparison, control comparison, or baseline comparison; see Table 4 and Figure 1). Each effect size was evaluated for its significance ($p \leq .05$) via the Comprehensive Meta-Analysis computer program (Version 1.25; Borenstein & Rothstein, 2002). Additionally, to control for escalating Type I error due to multiple statistical tests, the results are also reported with a Bonferroni Correction ($p \leq .001$). Table 4 and Figure 1 each reveal a number of significant effect sizes across a variety of cognitive domains and methods of evaluating change in patient performance post treatment. The results that remained significant with the conservative Bonferroni correction will be emphasized henceforth.

The results will be reviewed in several ways: (1) considering statistical significance, (2) considering absolute magnitude of effects, (3) considering relative magnitude of effects by examining patterns of findings within cognitive domains and methods of comparison, and (d) considering possible moderator variables.

Statistical Significance of Effects

It is noteworthy that 20 of 21 averaged weighted effect sizes across cognitive domains were in the negative direction, indicating a general trend toward decrements in functioning. Highly significant effects ($p \leq .001$) were found across five cognitive domains for two of the three methods of comparison (i.e., normative and control comparisons), while none of the seven cognitive domains were significant when the method of comparison was with one's own base-

Table 3. Unweighted effect sizes (Cohen's *d*) for individual studies by type of comparison and cognitive domain

First author	Year	Norm							Control							Baseline						
		A	*I*	R	*V*	S	*E*	M	A	*I*	R	*V*	S	*E*	M	A	*I*	R	*V*	S	*E*	M
Ahles (Breast)	2002	−0.15	*0.59*	−0.56	*0.53*	−0.25	*−0.28*	−0.75	−0.02	*−0.37*	−0.28	*−0.09*	−0.46	*−0.17*	0.00							
Ahles (Lymph)	2002	−0.12	*0.14*	−0.92	*0.14*	−0.31	*−0.32*	−0.75	−0.21	*−0.50*	−0.50	*−1.32*	−0.16	*−0.25*	−0.52							
Peace	2002								−0.26	*−0.79*	−0.62	*−0.34*	−0.31	*−0.30*								
Capuron	2001																*0.00*				*−1.39*	0.00
Riggs	2001								0.10	*−1.22*	0.48	*−0.75*	−0.20	*−0.29*	−0.07	0.90	*0.57*	0.17	*1.55*	−0.90	*1.35*	0.19
Brezden	2000								0.09		−0.19		−0.24	*−0.47*								
Mayerhofer	2000															0.00	*.79*					0.00
Schagen	1999								−0.49	*−0.32*	−0.27	*−0.48*	−0.24	*−0.29*	−0.44							
van Dam	1998								−0.14	*−0.27*	−0.30	*−0.41*	−0.25	*−0.18*	−0.33	0.06	*−0.76*	−0.40	*−0.69*	−0.57	*−0.05*	−0.69
Walker	1997								−1.61													
Ahles	1996			−0.19		−1.01											*−0.26*				*−0.51*	
van Oosterhout	1996								−0.57	*−0.97*	−1.06			*−0.66*								
Walker	1996													*−1.89*			*−3.14*					
Komaki	1995	−0.15	*−0.20*	−2.25	*−0.20*	0.30	*−0.55*	−0.65														
Meyers	1995	−0.20	*−0.01*	−1.91	*−0.15*	0.27	*−0.20*	0.11	0.31	*0.03*	−0.13	*0.02*	0.21	*0.07*	−0.02							
Pavol	1995	−.60	*−.30*	−1.85	*.20*	−.76	*−.70*	−.65	−0.90	*−0.70*	−0.65	*0.10*	−.70	*−1.23*	−0.70							
Wieneke	1995	−0.49	*−0.23*	−0.39	*−1.30*	−1.30	*−0.15*	−0.70														
Ahles	1994						*−2.40*														*−0.07*	
Archibald	1994						*−1.08*				−1.52	*−3.35*		*−2.40*								
Cull	1994	−2.82	*−0.77*				*−2.34*															
Meyers	1994															−0.16		0.10		0.17		
Andrykowski	1992	0.02	*0.00*	−0.36	*−0.28*		*−0.27*	−0.06														
Caraceni	1992															0.14	*−0.05*					
Meyers	1992	−0.45		−0.02	*−0.37*		*−0.88*															
Parth	1989								0.00	*−1.77*			−1.13	*−0.95*	0.00							
Tucker	1989			0.07	*0.53*																	
Kaasa	1988	−0.73					*−0.35*															
Denicoff	1987	0.16					*−1.13*		−0.65	*0.17*			0.18	*−0.22*		−0.66	*0.66*				*0.16*	−1.13
Oxman	1980	−1.00	*0.33*				*−1.18*															
Silberfarb	1980						*−2.15*															

Note. If more than one effect size could be extracted from a study in one of the seven cognitive domains, the average of those effect sizes was used. A = attention; I = information processing; R = verbal memory; V = visuospatial memory; S = visuospatial skill; E = executive function; M = psychomotor skill.
Important: Alternate columns are italicized to enhance readability only.

Table 4. Results for cognitive domains by method of comparison

Method	Cognitive domain	Attention	Info proc	Verb mem	Vis mem	Spat skill	Exec fn	Motor fn
Normative	*n* (studies)	12	9	**9**	10	6	**16**	7
	n (samples combined[1])	833	598	**532**	595	340	**1179**	**450**
	wt'd *d*	−.54	−.07	**−.91**	−.11	−.22	**−.93**	**−.48**
	99% CI–lower	−1.28	−.47	**−1.60**	−.53	−.83	**−1.11**	**−.85**
	99% CI–upper	.20	.32	**−.22**	.32	.38	**−.79**	**−.12**
	p value	.06	.63	**.001*^**	.51	.34	**<.001*^**	**.001*^**
	Fail-safe *N*	–	–	168	–	–	697	36
Control	*n* (studies)	**12**	**13**	**11**	**8**	**11**	**14**	**8**
	n (samples combined)	**702**	**675**	**615**	**439**	**634**	**751**	**424**
	wt'd *d*	**−.24**	**−.70**	**−.61**	**−.41**	**−.28**	**−.61**	**−.27**
	99% CI–lower	**−.49**	**−1.12**	**−1.22**	**−.82**	**−.55**	**−1.01**	**−.52**
	99% CI–upper	**.02**	**−.29**	**.00**	**.01**	**−.01**	**−.20**	**−.02**
	p value	**.02**	**<.001*^**	**.01^**	**.01^**	**.01^**	**<.001*^**	**.01^**
	Fail-safe *N*	16	188	87	25	22	154	7
Baseline	*n* (studies)	6	7	3	3	4	6	4
	n (samples combined)	154	147	68	81	119	154	80
	wt'd *d*	−.02	−.25	−.12	.16	−.24	−.28	−.20
	99% CI–lower	−.57	−1.40	−.62	−1.32	−.90	−1.16	−.81
	99% CI–upper	.53	.91	.38	1.63	.41	.60	.41
	p value	.93	.58	.53	.78	.34	.41	.39
	Fail-safe *N*	—	—	—	—	—	—	—

Note: *bold* figures represent significant results ($p \le .05$); * indicates a finding that is significant with a Bonferroni correction for multiple tests ($p \le .001$); ^ indicates a finding that remains significant even when limiting analyses to only studies of "less severe" treatments/diagnoses. [1]The *n*s for the "samples combined" for the normative comparisons exceed the total number of actual participants in the combined studies because these figures also include actual or estimated sample sizes of the normative comparison groups (a necessary figure for meta-analytic computations).

line. Thus, the findings appear more robust and consistent for the methods using between-subject comparisons as compared with studies that involved within-subject comparisons.

Absolute Magnitude of Effects

According to Cohen's (1988) guidelines (i.e., small effect = .2, medium effect = .5, large effect = .8), of the five highly significant effects, two effects exceeded the large effect size, while the remaining three approached or exceeded the medium effect size. The effect sizes for the normative comparisons of verbal memory and executive function were exceptionally large (both approximately .9).

Additionally, the number of studies needed to nullify the significant findings ("the file drawer statistic" or "fail-safe N") was calculated for each of these effect sizes (Rosenthal, 1991; see Table 4). Some of the effect sizes emerged as likely to remain significant even with a large number of additional studies yielding non-significant results. Rosenthal (1991) recommends a tolerance level for the file drawer statistic of $5K + 10$ (K = the number of studies used to calculate the effect size). Fail-safe Ns of 711 and 154 were calculated for the executive functioning effects for normative and control, respectively. This indicates the number of additional studies of each comparison type with null results that would be required for the significance of the reported finding to be reduced a non-significant level. The

fail-safe Ns for normative and control comparisons exceed Rosenthal's recommended cut-offs of 95 and 80 suggesting that these are very resilient findings unlikely to be disproven by a large number of studies yet unretrieved. Similarly, 168 (exceeds 55 cut-off) studies would be required to refute the verbal memory results for normative comparisons, while only 36 (does not exceed 45 cut-off) normative comparison studies would be necessary for refuting motor functioning findings.

Relative Magnitude of Effects Patterns Across Cognitive Domains and Methods of Comparison

Three cognitive domains revealed fairly consistent results across two methods of comparison. Effect sizes for executive function, verbal memory, and motor functioning were statistically significant (at either the .01 or .001 levels) whether the method of comparison was with normative data or control subjects. The effect sizes for executive function and verbal memory ranged from medium to large for both methods of comparison, while the effects for motor function ranged from small to large.

The baseline comparisons for these three domains were non-significant and the averaged weighted effect sizes were in the small to negligible range. It is possible that this is in part an artifact of the method of comparison since there are
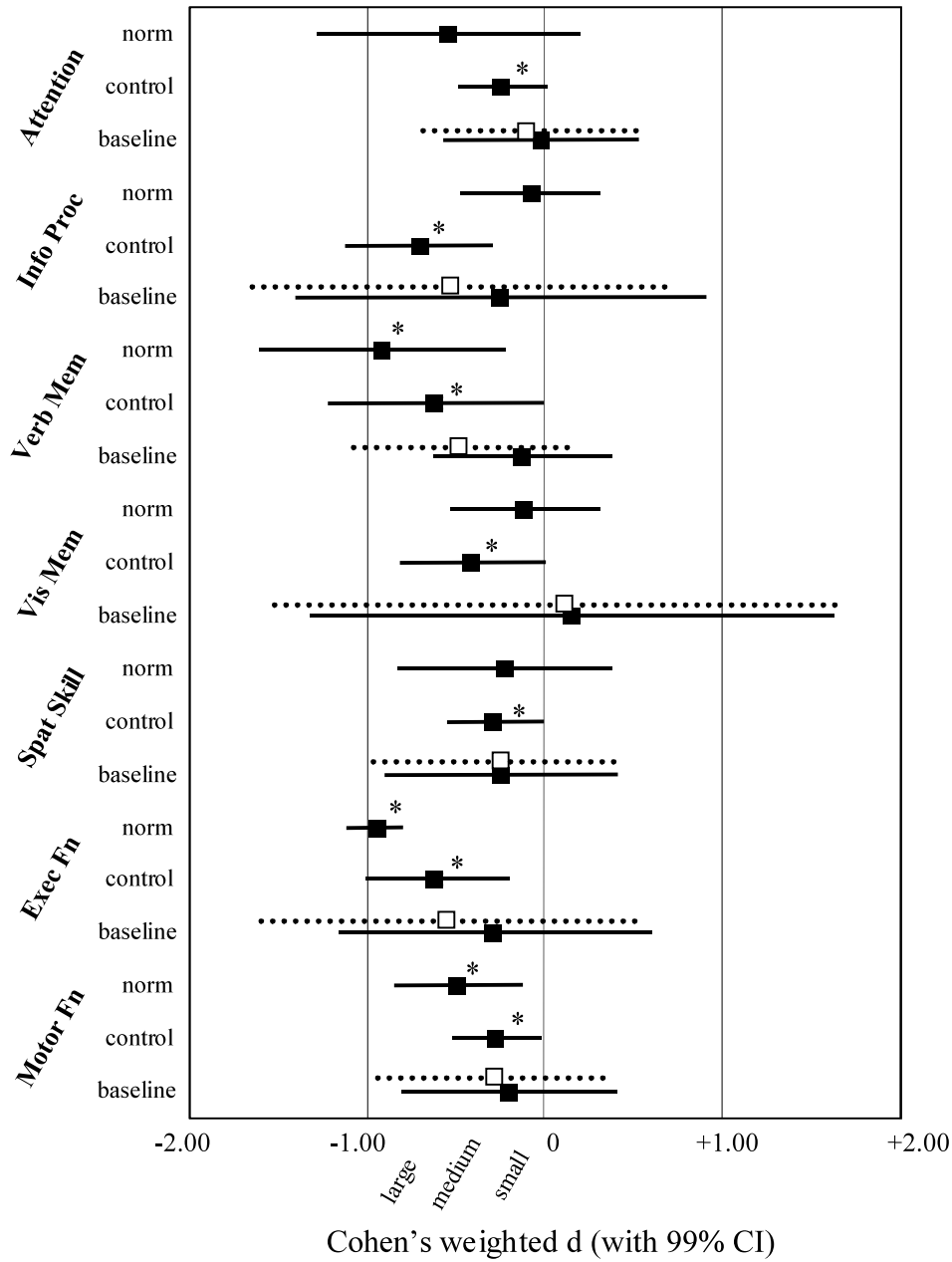
Fig. 1. Effect size results: Cognitive Domains × Method of Comparison. *$p \leq .05$. ··········□··········
indicates *d* and *CI* adjusted for practice/learning effects (Dikmen et al., 1999).

substantial learning and practice effects expected on some neuropsychological tasks when one's own baseline experience is compared with later retesting (even if alternate forms are used, which was not often the case in these studies; Dikmen et al., 1999). For a relative comparison of the baseline data, see the revised point estimates and dashed line confidence intervals in Figure 1 that represent an adjustment given the typical effect size found when using repeated measures (Dikmen et al., 1999; Temkin et al., 1999). A quick comparison reveals that the adjusted weighted effect sizes start to look more like the effect sizes of the normative and control comparisons than their unadjusted

counterparts. The relative differences of the unadjusted and adjusted effect sizes and confidence intervals suggest that perhaps a different mode of analysis is required when evaluating cognitive domain outcomes compared with one's own baseline. Rather than an absolute effect size contrasted with zero, a relative change from the expected normative practice/ learning curve effect could indicate a "significant" effect. For example, while the baseline comparison for executive function reveals an effect size of −.28 compared with zero (a small, non-significant effect), the adjusted effect size (taking into account the average learning/practice normative effect size on executive function tests) is approxi-

mately −.46, approaching a medium effect size (see adjusted point estimate and dashed confidence interval; CI, in Figure 1).

Because of the different types of control subjects utilized across studies (e.g., healthy controls *vs.* cancer patients with differing treatments), we re-analyzed these two groups of studies separately for each of the seven cognitive domains. This was done in an effort to rule out a possible bias (e.g., if most of the studies utilized healthy controls who might not have similar levels of fatigue or distress, this could artificially inflate the difference between cancer patients and these controls). Analyses revealed that five studies used healthy controls, while nine studies used cancer patients as controls. All seven comparisons with cancer controls remained statistically significant, even when omitting the health controls. In contrast comparisons with health controls sometimes became non-significant (with effect sizes shrinking) or sometimes became more highly significant (with effect sizes enlarging). In some domains there were only a handful of studies that utilized a health control group and this could have skewed results as well. It seems clear that even if studies are excluded that utilized health controls (which theoretically could artificially increase treatment effects due to the relative comparisons), the significant findings reported here across all seven cognitive domains are retained.

It is also possible to look for consistency, not only across cognitive domains, but also across methods of comparison. In applying this approach, it appears that normative and control comparisons yielded significant effects across several cognitive domains. In contrast, baseline comparisons yielded no significant effects; however, as noted above there are problems in interpreting this apparent lack of effect. The significance test determines if the effect size is significantly different than zero, but for baseline comparisons it may make more sense to compare any effect to the typically positive effect size that occurs with repeated measures (see above and dashed confidence intervals in Figure 1 which represent practice/learning effects found in normative samples due to repeated measures, even with alternate forms).

Possible Moderator Variable

Given the possible untoward effects of the heterogeneity of diagnoses and treatments on effect sizes, we recalculated effect sizes excluding those studies that included patients with more severe diagnoses or treatments (e.g., some patients within the study's sample may have had total brain irradiation, metastatic disease, CNS disease, or immunologic therapies). Twenty studies were deemed more severe and thus, nine studies were retained for this re-analysis (see Table 1). Table 4 reveals that 9 of the 10 statistically significant findings with the entire sample of studies remain significant even with only the subset of less severe diagnoses and treatments. Thus, it appears that the studies with more severe components were not having an untoward effect on the results.

DISCUSSION

The number of individuals living with or having survived cancer is increasing as a function of earlier disease detection and the development of more advanced treatments (National Cancer Institute, 1998). Given this reality, there is a growing need to understand the long-term impact of cancer treatments and their side effects on those who receive them. Recent research has begun to examine the impact of cancer treatments on short- and long-term quality of life, and some of these studies have explored the possibility that there are specific cognitive effects of cancer treatments such as chemotherapy. Reviews of the literature in this area have thus far been limited to strictly qualitative discussions (e.g., Ahles & Saykin, 2001; Olin, 2001). The intent of this meta-analysis was to provide a quantitative review that more clearly delineates the nature and magnitude of the effects detected in existing studies that have examined the impact of systemic cancer therapies on cognitive function. Additionally, it was hoped that the results of this quantitative review would guide future research in this area.

This meta-analysis encompassed 30 studies (leading to 29 samples) for which sufficient information was available to calculate an effect size for mean differences between the cognitive function of individuals who had received systemic treatments for cancer compared to either normative data, a control group, or their own pre-treatment baseline functioning. There are several important findings and implications of this meta-analysis.

When considering all identified studies, we found that for all averaged effects calculated across various cognitive domains and study designs, all significant findings were in the negative direction indicating that mean cognitive test scores for cancer patients having received systemic treatment were on average lower than those obtained from normative samples, study control groups, or pretreatment baseline assessment of the same cancer patients. The absolute magnitude of these effects across all studies ranged from negligible to large in size. The most consistent results were found across two methods of comparison (normative and control), for executive function, verbal memory and motor function[2]. The largest effect sizes were found for executive function and verbal memory when compared with normative data ($d = -.89$ and $-.91$, respectively).

Despite these consistencies across the normative and control comparison, the baseline comparison effects remain somewhat puzzling in that across the seven cognitive domains, these effect sizes remain smaller than their normative and control comparison counterparts. In part, this could be due to the fact that at the time of baseline testing cancer

---

[2]It is interesting to note that a study that was published just as this manuscript was going to press (Harder et al., 2002) reported results that are consistent with the results of this meta-analysis. Harder and colleagues similarly reported that impairments were found in patients years after bone marrow transplantation when compared with normative data on selective attention, executive function, information processing speed, verbal learning, and verbal and visual memory.

patients may already been experiencing some cognitive decline due to the effects of stress, fatigue or the sometimes toxic byproducts of the cancer process itself (e.g., pro-inflammatory cytokines). If so, later comparisons will not show much relative decline, potentially obscuring any real decline from premorbid functioning. Additionally, the absence of a significant effect in comparison with baseline functioning could be re-interpreted in light of normative data (e.g., test–retest), which typically shows positive effects/improvements in scores, due to practice and learning effects (Dikmen et al., 1999; McCaffrey et al., 2000; Salinsky et al., 2001; Temkin et al., 1999). In contrast, the effects for normative and control comparisons are interpreted relative to zero, which represents the mean for normative and control samples. If one instead considers the *relative change*, then the baseline comparison effect sizes for each cognitive domain appear more consistent with the effect sizes computed via normative and control comparisons (see effect sizes adjusted to account for this in Figure 1). Thus, the fact that six of seven cognitive domain effects as estimated by baseline comparisons are in the negative direction (*vs.* expected positive/increase), could be interpreted to mean that indeed treatments are having some undesirable effects.

In a more conservative analysis of only a subset of studies with the least intense treatments or least severe diagnoses (omitting any study with brain irradiation, metastatic disease, etc.), we found fairly consistent results. Nine of the 10 original statistically significant findings were retained, with the most consistent findings remaining (e.g., executive functioning, verbal memory, and motor functioning). Thus, it appears that even when considering only the least severe diagnoses and least intense treatments, there are still significant neuropsychological effects of cancer treatments in adults.

Other possible moderator variables besides intensity of treatment or severity of diagnosis could be examined in future research (e.g., simultaneous/contributing effects of other medicines such as anti-emetics, role of stress, depression, or fatigue). For example, newer hormonal treatments (e.g., Tamoxifen and Roloxifene) have led to preliminary examinations of their effects and Bender's (2001) review of hormonal treatments in breast cancer suggests detrimental effects. However, research regarding hormonal treatments in other populations suggests a possible protective effect (e.g., Paganini-Hill & Clark, 2001; Yaffe et al., 2001). Similarly, the role of certain medications given in conjunction with treatments (e.g., anti-fatigue medications, such as Herceptin), have been thought to cross the blood brain barrier and could also play a role in affecting cognition. Synergistic effects of combination treatments have also been noted (e.g., radiation combined with chemotherapy may intensify effects leading to increased cellular toxicities). Clearly, more careful research is needed that attempts to document, control for and parse out the potentially confounding effects of various components of multimodal treatments.

Comparisons of data from cancer patients after systemic treatment with normative or control sample data have pro-

duced the most sizable effects, while within-subject comparisons of baseline and post-treatment assessments have generally resulted in smaller averaged effects. This is an intriguing pattern of findings given that this latter study design could provide the most convincing evidence for the presence or absence of cognitive effects of any cancer treatment due to the designs' control of possibly confounding variables (e.g., diagnosis, education, premorbid cognitive functioning). While the use of normative and control groups for comparison provides useful information, it is the longitudinal within-subjects designs that allows for the control of individual differences in the cognitive performance of patients prior to treatment—although, as noted above, longitudinal results can also be clouded if the baseline performance is affected by tumor metabolites, stress of diagnosis, or other factors. Given these complexities, it is possible that normative and control comparisons may overestimate decline, while baseline comparisons may underestimate decline. Nonetheless, such prospective studies might be thought to be the best designs to assess the true effects of systemic treatments on cognitive function and thus one might be tempted to conclude that cancer treatments have little effect on cognitive functioning. However, such an interpretation would be premature for several reasons.

First, there are a number of statistical concerns relevant to the within subject studies. There were typically only two or three studies available per cognitive domain, and these studies typically had the smallest samples and the greatest heterogeneity in terms of diagnoses and treatment. In contrast, the cross-sectional studies have larger sample sizes and are more homogenous which could account for the relatively stronger pattern of results. Furthermore, the standard deviations used in calculating Cohen's *d* will tend to be smaller in the within-subject design and could artificially inflate the effect size. Thus, the reliability of the effect sizes reported in the within subject design is subject to further investigation. Second, as noted above, smaller effects observed in repeated-measures designs could be due, in part, to practice or learning effects that are artifacts of the instruments used (possible even with the use of alternative forms). This could certainly obscure actual detrimental effects of treatments being investigated. It appears that some accounting must be made to allow for the impact of practice and learning effects (Dikmen et al., 1999). Few investigators have assessed cognitive function across time using alternate forms of tests. Future research should at least aim to minimize these effects by using alternate forms, including healthy controls assessed at similar time intervals so as to measure practice effects, and statistically controlling for typical practice and learning effects via methods outlined in the literature (e.g., Temkin et al., 1999).

The effect sizes detected across all studies for executive function, verbal memory, and motor function indicate that, on average, individuals having received systemic cancer therapies perform $\frac{1}{3}$ to nearly 1 standard deviation below normative samples or a control groups in these domains. This level of disturbance is below that which would typi-

cally warrant a label of impairment by traditional neuropsychological standards (typical definitions range from -1 to -3 *SD*s). The deficits found here would not necessarily be expected to translate into easily observable functional difficulties in most patients, although given the variability in each sample there are likely some patients that do exhibit such profound and obvious problems. However, if these overall effect sizes are indeed accurate indications of the neuropsychological effects of such treatments, it is not surprising that they have been difficult to detect with small samples or clinical observation.

It is possible to better appreciate the level of deficit patients may be experiencing if we translate our findings into a more easily interpretable index, that is the percentage of people falling under any area of the normal curve. Mapping out the results in this way for the cognitive domain of executive function, we learn that while an average person who has not received any systemic cancer treatment would perform at the 50th percentile, the average individual having received such treatment might be expected to fall somewhere between the 15th and 35th percentile (i.e., their performance is below that of 65–85% of the normative comparison sample). Viewed in this way, it becomes more understandable why some patients may be more disturbed by and convinced of a deficit in cognitive functioning even when the scientific and medical communities struggle to verify (or disprove) its existence. Clearly an individual of average pre-morbid ability who can no longer perform cognitively as well as most of the individuals s/he once bested may experience this as distressing. For individuals of high pre-morbid ability who would perform at the upper end of the normal curve, the loss of $\frac{1}{3}$ to 1 standard deviation would translate into a smaller percentile difference. However, even such small differences may be experienced as detrimental to an individual who is used to performing in a professional position that demands peak cognitive skills.

A common theme that may tie these deficits across domains together could be an underlying phenomenon of slowing in mental and/or physical processing. Fatigue, which has been well documented as a frequent side effect of treatment that may not resolve readily with the end of treatment (Piper, 1998), may account for these effects. Anecdotally, patients often report that they just don't feel as "sharp" as they did pre-treatment or that their mind doesn't seem to "work as quickly" as it did in the past. Indeed, treatments for fatigue such as central nervous system stimulants (e.g., Ritalin) have been used to improve quality of life and apparently cognitive functioning in cancer patients (e.g., Weitzner et al., 1995). This suggests a possible underlying mechanism that might lie in subcortical processes or other pathways that could affect basic speed of neuronal functioning (as in motor functioning) and/or slow more complex higher order cognitive functioning seen in executive functioning, perhaps specifically a working memory component that may overlap with the verbal memory and executive function deficits found here. Fatigue is only one of many possible underlying mechanisms that could explain changes

in cognitive functioning. Furthermore, it is possible that there exists some "third variable" that leads to a syndrome or cluster of symptoms including both fatigue and cognitive problems, such as a common biological response to chemotherapy. Other changes in central nervous system functioning could be influential as well, such as white matter changes perhaps due to demylination, vascular changes, or formation of neurotoxic cytokines or free radicals that could cross the blood-brain barrier. Similarly, impacts of treatments on the peripheral nervous system (well documented in some cases) could account for decrements reported here, especially in motor functioning, but also for other cognitive domains. Clearly more research is needed to illuminate possible mechanisms.

Limitations

Taking into consideration that published data may be biased toward positive findings, with other null findings being present in unpublished data, this meta-analysis is subject to the file drawer problem described by Rosenthal (1991). However, calculations suggest largely robust findings for the control and normative comparisons. While every attempt was made to locate all available data, it is possible that other studies have been conducted for which data were not accessible. Such studies could be expected to support the null hypothesis (no difference between groups or assessment points) and would contribute to smaller average effect sizes than are reported here. Additionally, some studies had to be excluded due to a lack of sufficient statistical information in a report, most likely due to variability in journal publication requirements (e.g., the use of cut-offs or categorization of continuous variability *vs.* the use of means and variability measures). Future research would do well to promote more uniformity in reporting of results. Perhaps at a minimum, means and standard deviations for subgroups need to be included in all published studies in this area.

Another limitation of this meta-analysis is the nonspecificity with respect to a particular cancer diagnosis or treatment. There was a need to pool studies including various disease groups and treatment regimens given the paucity of data on this topic and the fact that studies did not always specify the diagnostic group(s) or specific systemic treatment regimen(s) delivered in the cancer patient sample. This led to the inclusion of a wide variety of diagnoses (e.g., breast, prostate, lung) and treatments (e.g., cyclophosphomide, interleukin-2, Tamoxifen) in these analyses. This heterogeneity can be viewed both as a weakness and a strength of this study. Heterogeneity can make interpretation of the analyses more difficult, but it can also point to the generality of the findings. Heterogeneity allows for a more rigorous test of the hypotheses; when there is a significant finding despite of heterogeneity, this suggests that the finding is quite resilient. This meta-analysis is meant to serve as a first, broad attempt to quantify the possible cognitive effects of cancer treatments and should not be used as an indicator of any specific treatment's neuropsychological

effects. Only with additional well-designed research in this area will it be possible to exact conclusions regarding a specific regimen or drug.

Strengths

In spite of the limitations noted, this meta-analysis provides a much-needed quantitative overview of the research on neuropsychological effects of cancer treatments. Rather than describing cognition as a general function, this meta-analysis isolated various cognitive domains and calculated separate effect sizes for each domain. In addition, separate effect-size analyses were conducted for the different study designs and subsequent methods of comparison used (i.e., between-subjects with normative data comparison, between-subjects with control group comparison, or within-subjects repeated measures comparison).

Furthermore, the control group comparison was explored in more detail since one could argue that the comparison of patients with matched healthy controls could lead to different results than comparison of patients with cancer patients who were receiving a different (a milder treatment such as localized radiation), or no treatment at all. It is possible to argue that the comparison with other cancer patients would be preferred, in that these patients would presumably be matched on other variables that could affect cognitive function (e.g., distress, fatigue). In fact, we found that even when considering cancer patient controls alone, analyses revealed the same significant results. Thus, we were reassured that the finding of a significant effect in the combined control sample is genuine (that is, the healthy controls are not unduly overpowering an otherwise null effect size).

By examining the effects in specific cognitive domains, the possible role of study designs and the variety of control subjects, this meta-analysis provides a more detailed picture of study results as they relate to both methodology and specific neuropsychological function.

Conclusions and Suggestions for Future Research

Undoubtedly, the strongest statement that can be made based on these analyses is that there is a great need for additional well-designed studies investigating this phenomenon. The data are consistent with the idea that there is an adverse effect, with the magnitude of the effect being more pronounced in some domains (e.g., executive function, verbal memory and motor function) than others. The strongest effect found in these analyses was based on a pooled sample of 511 patients. However, some other effects were based on fewer than 70 individuals. Clearly it is necessary to collect data from larger numbers of individuals in order to draw accurate and reliable inferences. In addition to data from a larger number of subjects, the studies conducted in this area must be designed to answer the question at hand: specifi-

cally, does a specific cancer treatment cause cognitive decline in those who receive it? To date, the data in this area consist primarily of a single assessment collected post-treatment and compared to a normative or control sample. Such cross-sectional between-subject designs are not fully adequate to determine whether any detected difference is truly due to the cancer treatment received rather than to some other individual difference such as pre-morbid cognitive ability. Through research using a longitudinal repeated-measures design, including a pre-treatment assessment, we will gain a better understanding of the causal nature of any effects seen. Future studies must incorporate such designs or the data will add little to what we already know. Ultimately, it would be best to see a clear convergence of results from all three types of study designs (methods of comparison). The needs for data from a larger pool of subjects as well as more within-subjects designs are both underscored by the fact that the effect sizes from the within-subject findings in this meta-analysis are the most tenuous.

Given that neuropsychological functioning is a multi-faceted phenomenon, future studies should continue to use test batteries that assess various cognitive domains. In this way, differential effects on specific cognitive abilities may be detected. The results of this meta-analysis suggest that executive function, verbal memory and motor function may be areas of particular interest to future investigators. However, other cognitive domains discussed above should not be ignored as these phenomena are still in the very early stages of investigation. Use of alternate test forms should also be part of the standard repeated-measures study design whenever reliable alternate forms of a test are available for a given cognitive domain. This will help minimize the washing out of any true differences due to practice or learning effects. Additionally, comparisons of results relative to typical practice or learning effects should be considered as well.

Future studies must also assess additional factors that may play a role in both cancer treatment and neuropsychological function. Specifically, depression, anxiety, sleep disturbance and fatigue are likely candidates that may result from a cancer diagnosis and treatment, and which have documented impacts on cognitive performance (Daly et al., 2001; Tiersky et al., 1997). However, assessment of these factors has not historically been standard practice in the studies reviewed. In the handful of studies that did measure and control for one or more of these factors, the conclusion was generally that results retained their significance despite consideration of such potential mediators or moderators. Twelve of the studies included in our analyses attempted to determine the impact of potentially relevant mediator or moderator variables (e.g., age, type of treatment, fatigue, mood; Brezden et al., 2000; Capuron et al., 2001; Denicoff et al., 1987; Meyers et al., 1995; Oxman & Silberfarb, 1980; Peace et al., 2002; Riggs et al., 2001; Schagen et al., 1999; Silberfarb et al., 1980; van Dam et al., 1998; Walker et al., 1996; Wieneke & Dienst, 1995). Eleven of these twelve studies found that the variables they measured did not re-

move or significantly account for the cognitive declines identified in their samples; only Oxman and Silberfarb (1980) found some significant impact of such variables. This suggests that the effects observed may not be easily explained away by other variables.

Nonetheless, future research should still strive to measure potential mediator or moderator variables since a larger sample of studies might yield sufficient power to detect potentially small effects. Ganz (1998) argues that treatment characteristics (e.g., a dose effect) may explain some equivocal results. Several reviews have pointed to a number of other possible contributing factors in the apparent cognitive decline (e.g., side effects of other medications, stress, etc.), as well as possible physiological explanations for the presence of some data showing cognitive declines (e.g., frontal lobe dysfunction seen on EEG with some treatments, a link between Tumor Necrosis Factor and white matter disease in animal models; Meyers, 2001; Silberfarb, 1983; Silberfarb & Oxman, 1988; Tope et al., 1993). Even one's subjective experience could be important to assess, perhaps implicating a type of self-fulfilling prophecy (Cull et al., 1996; Devlen et al., 1987a, 1987b). Future research should incorporate these types of variables into sophisticated analyses that might detect possible relationships to individual neuropsychological performance. Investigators must be mindful to assess additional variables, such as those suggested above, that may impact the effects being explored and use appropriate analytical techniques to examine those variables as mediators, moderators, or confounds in the relationships identified.

In summary, given the data available to date, it appears that cancer treatments in adults can impact neuropsychological functioning in those who receive them. These effects fall primarily in the mild to moderate range and warrant further investigation. The most consistent effects across most methods of comparison are seen in executive function, verbal memory and motor skills. The implications of these findings for the clinician are not yet clear. As the first quantitative review, caution should be used in interpreting this data and future research should lead to updated meta-analyses that might clarify the finer points of interest to clinicians and their patients. Future research may clarify the possible need for a more detailed informed consent process for cancer treatment and more focused quality of life discussions between clinicians and their patients. Future research may provide some validation of some patients' experiences, it might increase anxiety for others anticipating treatment, or relieve those who can anticipate smaller or more reversible declines than they might imagine. Additionally, interventions might be designed to assist patients in coping with any anticipated cognitive changes (e.g., Ferguson & Ahles, 2002). Clearly, it would be immensely helpful if we could clarify which specific treatments might be expected to lead to a detrimental effect and how long any specific effect might last so that patients can make informed choices and prepare for any impact on work or relationships. This meta-analysis serves to highlight some

aspects of a complex jigsaw puzzle that still has many pieces missing.

REFERENCES

Ahles, T.A. & Saykin, A.J. (2001). Cognitive effects of standard-dose chemotherapy in patients with cancer. *Cancer Investigator*, *19*, 812–820.

*Ahles, T.A., Saykin, A.J., Furstenberb, C.T., Cole, B., Mott, L.A., Skalla, K., Whedon, M.B., Bivens, S., Mitchell, T., Greenberg, R., & Silberfarb, P.M. (2002). Neuropsychological impact of standard-dose systemic chemotherapy in long-term survivors of breast cancer and lymphoma. *Journal of Clinical Oncology*, *20*, 485–493.

*Ahles, T.A., Silberfarb, P.M., Rundle, A.C., Holland, J.C., Kornblith, A.B., Canellos, G.P., Green, M.R., & Perry, M.C. (1994). *Psychotherapy Psychosomatics*, *62*, 193–199.

*Ahles, T.A., Tope, D.M., Furstenberg, C., Hann, D., & Mills, L. (1996). Psychologic and neuropsychologic impact of autologous bone marrow transplantation. *Journal of Clinical Oncology*, *14*, 1457–1462.

*Andrykowski, M.A., Schmitt, F.A., Gregg, M.E., Brady, M.J., Lamb, D.G., & Henlsee-Downey, P.J. (1992). Neuropsychological impairment in adult bone marrow transplant candidates. *Cancer*, *70*, 2288–2297.

*Archibald, Y., Lunn, D., Ruttan, L., Macdonald, D., Del Maestro, R., Barr, H., Pexman, W., Fisher, B., Gaspar, L., & Cairncross, G. (1994). Cognitive functioning in long-term survivors of high-grade glioma. *Journal of Neurosurgery*, *80*, 247–253.

Bender, C.M., Paraska, K.K., Sereika, S.M., Ryan, C.M., & Berga, S.L. (2001). Cognitive Function and Reproductive Hormones in Adjuvant Therapy for Breast Cancer: A Critical Review. *Journal of Pain and Symptom Management*, *21*, 407–424.

Bonadonna, G., Valagussa, V.E., & Rossi, A. (1985). Ten-year experience with CMF-based adjuvant chemotherapy in resectable breast cancer. *Breast Cancer Research and Treatment*, *5*, 95.

Borenstein, M. & Rothstein, H. (2002). *Comprehensive Meta Analysis Computer Program and Manual*. Englewood Cliffs, NJ: Biostat, Inc.

*Brezden, C.B., Phillips, K., Abdolell, M., Bunston, T., & Tannock, I.F. (2000). Cognitive function in breast cancer patients receiving adjuvant chemotherapy. *Journal of Clinical Oncology*, *18*, 2695–2701.

*Indicates that this study was included in analyses.

Brundage, M.D. (1997). Trading treatment toxicity for survival in locally advanced non-small cell lung cancer. *Journal of Clinical Oncology*, *15*, 330–340.

Butler, R.W. (2002). Attentional processes and their remediation in children treated for cancer: a literature review and the development of a therapeutic approach. *Journal of the International Neuropsychological Society*, *8*, 115–124.

*Capuron, L., Ravaud, A., & Dantzer, R. (2001). Timing and specificity of the cognitive changes induced by Interleukin-2 and Interferon-*α* treatments in cancer patients. *Psychosomatic Medicine*, *63*, 376–386.

*Caraceni, A., Martini, C., Belli, F., Macheroni, L., Rivoltini, L., Arienti, F., & Cascinelli, N. (1992). Neuropsychological and neurophysiological assessment of the central effects of interleukin-2 administration. *European Journal of Cancer*, *29A*, 1266–1269.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooper, H. & Hedges, L.V. (1994). *Handbook of research synthesis*. New York: Russell Sage Foundation.

Copeland, D.R., Fletcher, J.M., & Pfefferbaum-Levine, B. (1985). Neuropsychological sequelae of childhood cancer in long-term survivors. *Pediatrics*, *75*, 745–753.

*Cull, A., Gregor, A., Hopwood, P., Macbeth, F., Karnicka-Mlodkowska, H., Thatcher, N., Burt, P., Stepniewska, K., & Stewart, M. (1994). Neurological and cognitive impairment in long-term survivors of small cell lung cancer. *European Journal of Cancer*, *30a*, 1067–1074.

Cull, A., Hay, C., Love, S.B., Mackie, M., Smets, E., & Stewart, M. (1996). What do cancer patients mean when they complain of concentration and memory problems? *British Journal of Cancer*, *74*, 1674–1679.

Daly, E., Komaroff, A.L., Bloomingdale, K., Wilson, S., & Albert, M.S. (2001). Neuropsychological function in patients with chronic fatigue syndrome, multiple sclerosis, and depression. *Applied Neuropsychology*, *8*, 12–22.

*Denicoff, K.D., Rubinow, D.R., Papa, M.Z., Simpson, C., Seipp, C.A., Lotze, M.T., Change, A.E., Rosenstein, D., & Rosenberg, S.A. (1987). The neuropsychiatric effects of treatment with interleukin-2 and lymphokine-activated killer cells. *Annals of Internal Medicine*, *107*, 293–300.

Devlen, J., Maguire, P., Phillips, P., & Crowther, D. (1987a). Psychological problems associated with diagnosis and treatment of lymphomas: II. Prospective study. *British Journal of Medicine*, *295*, 955–957.

Devlen, J., Maguire, P., Phillips, P., Crowther, D., & Chambers, H. (1987b). Psychological problems associated with diagnosis and treatment of lymphomas. I: Retrospective study. *British Journal of Medicine*, *295*, 953–954.

Dikmen, S.S., Heaton, R.K., Grant, I., & Temkin, N.R. (1999). Test–retest reliability and practice effects of Expanded Halstead-Reitan Neuropsychological Test Battery. *Journal of the International Neuropsychological Society*, *5*, 346–356.

Eiser, C.J. (1998). Practitioner review: Long-term consequences of childhood cancer. *Child Psychology and Psychiatry*, *39*, 621–633.

Ferguson, R.J. & Ahles, T.A. (2002). *Brief cognitive-behavioral treatment of chemotherapy-associated attention and memory complaints among breast cancer survivors: Pilot data*. Poster presented at the Annual Meeting of the Society of Behavioral Medicine, Washington, DC.

Fisher, B., Costantino, J.P., & Redmond, C.K. (1994). Endometrial cancer in tamoxifen-treated breast cancer patients: Findings from the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-14. *Journal of the National Cancer Institute*, *86*, 527.

Fleishman, S.B. & Kalash, G.R. (1998). Chemotherapeutic agents and neuropsychiatric side effects. In J.C. Holland (Ed.), *Psychooncology* (pp. 630–658). New York: Oxford University Press.

Ganz, P.A. (1998). Cognitive dysfunction following adjuvant treatment of breast cancer: A new dose-limiting toxic effect? *Journal of the National Cancer Institute*, *90*, 182–183.

Gotay, C.C. (1987). Quality of life among survivors of childhood cancer: A critical review and implications for intervention. *Journal of Psychosocial Oncology*, *5*, 5–23.

Harder, H., Cornelissen, J.J., Van Gool, A.R., Duivenvoorden, H.J., Eijkenboom, W.M., & van den Bent, M.J. (2002). Cognitive functioning and quality of life in long-term adult survivors of bone marrow transplantation. *Cancer*, *95*, 183–192.

*Kaasa, S., Olsnes, B.T., & Mastekaasa, A. (1988a). Neuropsychological evaluation of patients with inoperable non-small cell lung cancer treated with combination chemotherapy or radiotherapy. *Acta Oncologica*, *27*, 241–246.

*Kaasa, S., Olsnes, B.T., Thorud, E., & Host, H. (1988b). Reduced short-term neuropsychological performance in patients with nonsmall-cell lung cancer treated with cisplatin and etoposide. *Antibiotics and Chemotherapy*, *41*, 226–231.

*Komaki, R., Meyers, C.A., Dong, M.S., Garden, A.S., Byrne, K., Nickens, J.A., & Cox, J.D. (1995). Evaluation of cognitive function in patients with small cell lung cancer prior to and shortly following prophylactic cranial irradiation. *International Journal of Radiation Oncology, Biology, and Physiology*, *33*, 179–182.

Mann, D. (1999). Memory lock: Is chemobrain making you forgetful? *MAMM: Women, Cancer and Community*, December/January, 67–69.

*Mayerhofer, K., Bodner-Adler, B., Bodner, K., Saletu, B., Schindl, M., Kaider, A., Hefler, L., Leodolter, S., & Kainz, C. (2000). A Paclitaxel-containing chemotherapy does not cause central nervous adverse effects: A prospective study in patients with ovarian cancer. *Anticancer Research*, *20*, 4051–4056.

McCaffrey, R.J., Duff, K., & Westervelt, H.J. (2000). *Practitioner's guide to evaluating change with intellectual assessment instruments*. New York: Kluwer Academic.

Meyers, C.A. (2000). Neurocognitive dysfunction in cancer patients. *Oncology*, *14*, 75–79.

Meyers, C.A. (2001). Mental status changes associated with biotherapy. In P.T. Rieger (Ed.), *Biotherapy: A comprehensive overview* (2nd ed., pp. 605–621). Boston: Jones and Bartlett Publishers.

*Meyers, C.A. & Abbruzzese, J.L. (1992). Cognitive functioning in cancer patients: Effect of previous treatment. *Neurology*, *42*, 434–436.

*Meyers, C.A., Byrne, K., & Komaki, R. (1995). Cognitive deficits in patients with small cell lung cancer before and after chemotherapy. *Lung Cancer*, *12*, 231–235.

Meyers, C.A. & Scheibel, R.S. (1990). Early detection and diagnosis of neurobehavioral disorders associated with cancer and its treatment. *Oncology*, *4*, 115–122.

*Meyers, C.A., Weitzner, M., Byrne, K., Valentine, A., Champlin, R.E., & Przepiorka, D. (1994). Evaluation of the neurobehavioral functioning of patients before, during and after bone marrow transplantation. *Journal of Clinical Oncology*, *12*, 820–826.

National Cancer Institute (1998). Surveillance, Epidemiology, and End Results (SEER) Program public-use data (1973–1998), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2001, based on the August 2000 submission. Bethesda, MD: Author.

Olin, J.J. (2001). Cognitive function after systemic therapy for breast cancer. *Oncology*, *15*, 613– 624.

*Oxman, T.E. & Silberfarb, P.M. (1980). Serial cognitive testing in cancer patients receiving chemotherapy. *American Journal of Psychiatry*, *137*, 1263–1265.

Paganini-Hill, A. & Clark, L.J. (2000). Preliminary assessment of cognitive function in breast cancer patients treated with Tamoxifen. *Breast Cancer Research and Treatment*, *64*, 165–176.

*Parth, P., Dunlap, W.P., Kennedy, R.S., Ordy, J.M., & Lane, N.E. (1989). Motor and cognitive testing of bone marrow transplant patients after chemoradiotherapy. *Perceptual Motor Skills*, *68*, 1227–1241.

*Pavol, M.A., Meyers, C.A., Rexer, J.L., Valentine, A.D., Mattis, P.J., & Talpaz, M. (1995). Pattern of neurobehavioral deficits associated with interferon-alfa therapy for leukemia. *Neurology*, *45*, 947–950.

*Peace, K., Freeman, J., Perez, D., & Tripp, G. (2002). *Long-term effects of adjuvant chemotherapy on cognitive functioning in patients with breast or bowel cancer: Preliminary findings*. Poster presented at the Annual Meeting of the International Neuropsychological Society; Stockholm, Sweden.

Peterson, L.G. & Popkin, M.K. (1980). Neuropsychiatric effects of chemotherapeutic agents for cancer. *Psychosomatics*, *21*, 141–153.

Petros, W.P. (2002). Association of high-dose cyclophosphamide, cisplatin, and carmustine pharmacokinetics with survival, toxicity, and dosing weight in patients with primary breast cancer. *Clinical Cancer Research*, *8*, 698–705.

Piper, B.F. (1998). The revised Piper Fatigue Scale: psychometric evaluation in women with breast cancer. *Oncology Nursing Forum*, 25, 677–684.

Redd, W.H., Silberfarb, P.M., Anderson, B.L., Andrykowski, M.A., Bovbjerg, D.H., Burish, T.G., Carpenter, P.J., Cleeland, C., Dolgin, M., & Levy, S.M. (1991). Physiological and psychobehavioral research in oncology. *Cancer*, *67*(suppl.), 813–822.

*Riggs, R., Sherman, M., Muss, H., Anderson-Hanley, C., O'Brien, P., & Compas, B. (May, 2001). A Prospective Study of the Neuropsychological Effects of Systemic Treatments for Early Stage Breast Cancer. Abstract published in the *Proceedings of the American Society for Clinical Oncology*, *20*, 29b.

Roman, D.D. & Sperduto, P.W. (1995). Neuropsychological effects of cranial radiation: Current knowledge and future directions. *International Journal of Radiation Oncology, Biology and Physics*, *31*, 983–998.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Rev. ed.). Newbury Park, CA: Sage.

Rosenthal, R. & DiMatteo, M.R. (2001). Meta-Analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, *52*, 59–82.

Salinsky, M.C., Storzbach, D., Dodrill, C.B., & Binder, L.M. (2001). Test–retest bias, reliability, and regression equations for neuropsychological measures repeated over a 12–16-week period. *Journal of the International Neuropsychological Society*, *7*, 597– 605.

Saphner, T., Tormey, D.C., & Gray, R. (1991). Venous and arterial thrombosis in patients who received adjuvant therapy for breast cancer. *Journal of Clinical Oncology*, *9*, 286–294.

*Schagen, S.B., van Dam, S.B., Muller, M.J., Boogerd, W., Lindeboom, J., & Bruning, P.F. (1999). Cognitive deficits after postoperative adjuvant chemotherapy for breast carcinoma. *Cancer*, *85*, 640– 650.

Shadish, W.R. & Haddock, C.K. (1994). Combining estimates of effect size. In H. Cooper & L.V. Hedges, (Eds.), *Handbook of Research Synthesis*, (pp. 261–281). New York: Russell Sage Foundation.

Silberfarb, P.M. (1983). Chemotherapy and cognitive defects in cancer patients. *Annual Review of Medicine*, *34*, 35– 46.

Silberfarb, P.M. & Oxman, T.E. (1988). The effects of cancer therapies on the central nervous system. *Advances in Psychosomatic Medicine*, *18*, 13–25.

*Silberfarb, P.M., Philibert, D., & Levine, P.M. (1980). Psychosocial aspects of neoplastic disease, II: Affective and cognitive effects of chemotherapy in cancer patients. *American Journal of Psychiatry*, *137*, 597– 601.

Temkin, N.R., Heaton, R.K., Grant, I., & Dikmen, S.S. (1999). Detecting significant change in neuropsychological test performance: A comparison of four models. *Journal of International Neuropsychological Society*, *5*, 357–369.

Tiersky, L.A., Johnson, S.K., Lange, G., Natelson, B.H., & DeLuca, J. (1997). Neuropsychology of chronic fatigue syndrome: A critical review. *Journal of Clinical and Experimental Neuropsychology*, *19*, 560–586.

Tope, D.M., Ahles, T.A., & Silberfarb, P.M. (1993). Psychooncology: Psychological well-being as one component of quality of life. *Psychotherapy Psychosomatics*, *60*, 129–147.

Trask, P.C., Esper, P., Riba, M., & Redman, B. (2000). Psychiatric side effects of interferon therapy: Prevalence, proposed mechanisms, and future directions. *Journal of Clinical Oncology*, *18*, 2316–2326.

Troy, L., McFarland, K., Littman-Power, S., Kelly, B.J., Walpole, E.T., Wyld, D., & Thompson, D. (2000). Cisplatin-based therapy: A neurological and neuropsychological review. *Psycho-Oncology*, *9*, 29–39.

*Tucker, J., Prior, P.F., Green, C.R., Ede, G.M., Stevenson, J.F., Gawler, J., Jamal, G.A., Charlesworth, M., Thakkar, C.M., Patel. P., & Lister, T.A. (1989). Minimal neuropsychological sequelae following prophylactic treatment of the central nervous system in adult leukemia and lymphoma. *British Journal Cancer*, *60*, 775–780.

*van Dam, F.S.A.M., Schagen, S.B., Muller, M.J., Boogerd, W., Wall, E., Fortuyn, M.E.D., & Rodenhuis, S. (1998). Impairment of cognitive function in women receiving adjuvant treatment for high-risk breast cancer: High-dose versus standard-dose chemotherapy. *Journal of the National Cancer Institute*, *90*, 210–218.

*van Oosterhout, A.G., Ganzevles, P.G., Wilmink, J.T., De Geus, B.W., Van Vonderen, R.G., & Twijnstra, A. (1996). Sequelae in long-term survivors of small cell lung cancer. *International Journal of Oncology, Biology and Physics*, *34*, 1037–1044.

Videtic, G.M. (2001). Using treatment interruptions to palliate the toxicity from concurrent chemoradiation for limited small cell lung cancer decreases survival and disease control. *Lung Cancer*, *33*, 249–258.

Walch, S.E., Ahles, T.A., & Saykin, A.J. (1998). Neuropsychological impact of cancer and cancer treatments. In J.C. Holland (Ed.), *Psycho-oncology*, (pp. 500–508). New York: Oxford University Press.

*Walker, L.G., Walker, M.B., Heys, S.D., Lolley, J., Wesnes, K., & Eremin, O. (1997). The psychological and psychiatric ef-

fects of rIL-2 therapy: A controlled clinical trial. *Psycho-oncology*, *6*, 290–301.

*Walker, L.G., Wesnes, K.P., Heys, S.D., Walker, M.B., Lolley, J., & Eremin, O. (1996). The cognitive effects of recombinant Interleukin-2 (rIL-2) therapy: A controlled clinical trial using computerized assessments. *European Journal of Cancer*, *32A*, 2275–2283.

Weiss, H.D., Walker, M.D., & Wiernik, P.H. (1974a). Neurotoxicity of commonly used antineoplastic agents (first of two parts). *New England Journal of Medicine*, *291*, 75–81.

Weiss, H.D., Walker, M.D., & Wiernik, P.H. (1974b). Neurotoxicity of commonly used antineoplastic agents (second of two parts). *New England Journal of Medicine*, *291*, 127–133.

Weitzner, M.A. & Meyers, C.A. (1997). Cognitive functioning and quality of life in malignant glioma patients: A review of the literature. *Psycho-Oncology*, *6*, 169–177.

Weitzner, M.A., Meyers, C.A., & Valentine, A.D. (1995). Methylphenidate in the treatment of neurobehavioral slowing associated with cancer and cancer treatment. *Journal of Neuropsychiatry and Clinical Neurosciences*, *7*, 347–350.

*Wieneke, M.H. & Dienst, E.R. (1995). Neuropsychological assessment of cognitive functioning following chemotherapy for breast cancer. *Psycho-Oncology*, *4*, 61–66.

Yaffe, K., Krueger, K., Sarkar, S., Grady, D., Barrett-Connor, E., Cox, D.A., & Nickelsen, T. (2001). Cognitive function in postmenopausal women treated with Raloxifene. *New England Journal of Medicine*, *344*, 1207–1213.