JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Longitudinal Assessment of Neurocognitive Outcomes in Survivors of Childhood Acute Lymphoblastic Leukemia Treated on a Contemporary Chemotherapy Protocol

Lisa M. Jacola, Kevin R. Krull, Ching-Hon Pui, Deqing Pei, Cheng Cheng, Wilburn E. Reddick, and Heather M. Conklin

A B S T R A C T

Purpose
Survivors of childhood acute lymphoblastic leukemia (ALL) treated with CNS-directed chemotherapy are at risk for neurocognitive deficits. Prospective longitudinal studies are needed to clarify the neurodevelopmental trajectory in this vulnerable population.

Methods
Patients enrolled in the St. Jude Total Therapy Study XV, which omitted prophylactic cranial radiation therapy in all patients, completed comprehensive neuropsychological assessments at induction (n = 142), end of maintenance (n = 243), and 2 years after completion of therapy (n = 211). We report on longitudinal change in neurocognitive function and predictors of neurocognitive outcomes 2 years after completing therapy.

Results
Neurocognitive function was largely age appropriate 2 years after completing therapy; however, the overall group demonstrated significant attention deficits and a significantly greater frequency of learning problems as compared with national normative data (all $P \leq .005$). Higher-intensity CNS-directed chemotherapy conferred elevated risk for difficulties in attention, processing speed, and academics (all $P \leq .01$). The rate and direction of change in performance and caregiver-reported attention difficulties differed significantly by age at diagnosis and sex. End-of-therapy attention problems predicted lower academic scores 2 years later, with small to moderate effect sizes ($|r| = 0.17$ to 0.25, all $P \leq .05$).

Conclusion
Two years after chemotherapy-only treatment, neurocognitive function is largely age appropriate. Nonetheless, survivors remain at elevated risk for attention problems that impact real-world functioning. Attention problems at the end of therapy predicted decreased academics 2 years later, suggesting an amplified functional impact of discrete neurocognitive difficulties. Age at diagnosis and patient sex may alter neurocognitive development in survivors of childhood ALL treated with chemotherapy-only protocols.

J Clin Oncol 34:1239-1247. © 2016 by American Society of Clinical Oncology

INTRODUCTION

Acute lymphoblastic leukemia (ALL) is the most common childhood malignancy, accounting for 29% of childhood cancer diagnoses.[1] Prognosis has improved substantially over several decades, with current 10-year overall survival rates reaching 90%.[2] Improved survival is attributed to therapeutic advances, including prophylactic CNS-directed treatment.[2] Contemporary therapy has largely replaced cranial radiation therapy (CRT) in favor of intrathecal (IT) chemotherapy for CNS prophylaxis. When compared with those treated with CRT, survivors treated with chemotherapy only show relatively preserved neurocognitive function[3,4]; however, the majority of studies suggest that these survivors continue to be at increased risk for deficits in attention,[5] processing speed,[5,6] and executive functions.[5,6] These deficits have been identified relative to healthy control participants, siblings, and age-based normative expectations, and negatively impact functional outcomes[7] and quality of life.[8] Greater intensity of CNS-directed chemotherapy[9,10] and younger age at diagnosis[11-13] have been most consistently identified as risk factors, with some evidence that female sex[13,14] is also a risk factor. Neurocognitive performance has been associated with treatment-related changes in brain structure and function.[15-17]

Few longitudinal studies have examined neurocognitive outcomes from active therapy to survivorship in patients treated with chemotherapy only. Harila et al[18] documented significant decline in verbal intelligence in 12 survivors who completed assessments at the end of treatment, 5 years post treatment, and 10 to 32 years post treatment. When compared with healthy control participants, survivors treated on the low-risk arm of UK Pediatric ALL Trial (UKALL) XI evidenced significant declines in global intelligence at assessment completed 3 to 5 years postdiagnosis.[19] A study of 17 survivors treated with chemotherapy only showed no significant differences on testing conducted at the end of therapy between patients and age-matched control participants; however, patients performed significantly worse than control participants on measures of attention and learning at a median follow-up of 5 years later.[12] A study of 49 patients tested at diagnosis, end of therapy, and 5 years later found that neurocognitive functioning was not significantly different than sibling control participants, with the exception of deficits on a measure of visual motor function in the patient group at the final assessment. However, a subgroup of 13 patients defined by elevated physical complaints at diagnosis had significantly lower attention and processing speed than siblings at the end of therapy and 5-year follow-up.[20] These results provide preliminary evidence of neurocognitive decline in survivors of ALL treated without CRT; however, these findings are limited by retrospective design[18] and small sample size.[12,18-20] Prospective longitudinal studies with large and representative cohorts are needed to clarify the neurocognitive trajectory after treatment, quantify the functional impact of these difficulties, and target areas for intervention.

The St. Jude Total XV Therapy Protocol (Total XV), which involved an intensification of systemic chemotherapy and optimal IT chemotherapy, allowed for complete omission of CRT without compromising overall survival.[21] This clinical trial included serial neurocognitive assessment with a comprehensive test battery. At the end of therapy, survivors were found to have a significantly increased risk for sustained attention deficits when compared with normative expectations.[10] When compared with survivors treated with lower-intensity therapy, patients treated with higher-intensity therapy had significantly reduced processing speed and academic scores and were at significantly greater risk for learning and behavior problems.

The current study examines the pattern of neurocognitive outcomes from end of therapy to 2 years after completion of Total XV. First, we hypothesize the overall group will demonstrate persistently elevated risk for attention problems 2 years after completing therapy. Second, we hypothesize a greater risk for neurocognitive deficits in survivors treated with greater-intensity therapy and those treated at a younger age. Given the association between attention problems and reduced academics previously demonstrated in survivors of childhood ALL,[7] we hypothesize that end-of-therapy attention will predict reduced academics 2 years after completion of therapy.

METHODS

Patients

A total of 408 patients 1 to 18 years old at diagnosis were consecutively enrolled on Total XV at St. Jude Children's Research Hospital between 2000 and 2007. Children were excluded from cognitive assessment if they had Down syndrome (n = 10) or did not speak English as a primary language (n = 16). Patients missing psychological data (children with refractory or progressive disease, refusals, missed appointments and scheduling difficulties; numbers varied by time point) were excluded from the analysis. The study was approved by the Institutional Review Board at St. Jude Children's Research Hospital. Written informed consent, with assent from the patient as appropriate, was obtained before assessment.

Treatment

Treatment details and primary outcomes have been previously described.[21] Children were assigned to low-risk or combined standard/high-risk groups on the basis of comprehensive biologic and clinical information, including blast cell immunophenotype and genotype, presenting clinical features, and early treatment response. All patients received triple IT chemotherapy with methotrexate (MTX), cytarabine, and hydrocortisone as CNS-directed therapy (dose ranges: low risk, 13 to 18; standard/high risk, 16 to 25) beginning with remission induction. During consolidation, patients were given four cycles of intravenous high-dose MTX (low risk: 2.5 g/m^2; standard/high risk: 5.0 g/m^2), titrated such that patients received the same systemic exposure of MTX (low risk: 33 μM;

	Table 1. Demographic and Clinical Characteristics			
	2 Years Post (N = 211)		End of Therapy + 2 Years Post (N = 167)	
	n (%)	Mean ± SD	n (%)	Mean ±SD
Sex				
Male	107 (50.7)		83 (49.7)	
Female	104 (49.3)		84 (50.3)	
Race/ethnicity				
White	170 (80.6)		131 (78.4)	
Black	33 (15.6)		29 (17.4)	
Other	8 (3.8)		7 (4.2)	
Age at diagnosis		6.49 ± 4.26		6.59 ± 4.25
< 5 years old	102 (48.3)		83 (49.7)	
≥ 5 years old	109 (51.7)		84 (50.3)	
Treatment arm				
Low risk	115 (54.5)		89 (53.3)	
Standard/high risk	96 (45.5)		78 (46.7)	
Induction Wechsler EIQ	102(48.3)	101.57 ± 16.52	88 (52.7)	102.25 ± 16.04

Abbreviation: EIQ, estimated IQ.

Table 2. Task Performance by Age at Diagnosis, Sex, and Risk Arm at 2 Years Post

	Age at Diagnosis					Treatment Arm					Sex				
	< 5 Years		≥ 5 Years			Low		Standard/High			Male		Female		
	Mean ± SD	P*	Mean ± SD	P*	P†	Mean ± SD	P*	Mean ± SD	P*	P†	Mean ± SD	P*	Mean ± SD	P*	P†
Sustained attention (CPT)															
Omissions	82.2 ± 20.8	‡	79.0 ± 20.7	‡	—	80.3 ± 19.8	‡	80.6 ± 21.8	‡	—	78.7 ± 23.2	‡	82.4 ± 17.4	‡	—
Attentiveness (D′)	60.2 ± 9.4	‡	57.3 ± 10.9	‡	—	58.9 ± 9.4	‡	58.3 ± 11.4	‡	—	58.0 ± 10.0	‡	59.3 ± 10.7	‡	—
Hit reaction time	51.7 ± 12.5	—	51.5 ± 13.3	—	‖	52.8 ± 12.0	‖	50.1 ± 13.8	—	—	51.7 ± 13.2	—	51.4 ± 12.6	—	—
Variability	60.6 ± 11.8	‡	56.4 ± 14.3	‡	—	58.7 ± 13.3	‡	57.8 ± 13.5	‡	—	56.5 ± 12.5	‡	60.3 ± 14.1	‡	—
β	70.3 ± 19.4	‡	70.0 ± 19.7	‡	—	68.6 ± 19.0	‡	72.0 ± 20.1	‡	—	70.2 ± 19.7	‡	70.2 ± 19.3	‡	—
Caregiver report (CPRS)															
Learning	54.7 ± 17.1	‖	50.7 ± 13.1	—	—	52.5 ± 16.3	—	53.5 ± 14.5	—	—	51.3 ± 13.2	—	54.5 ± 17.5	‖	—
Hyperactivity	50.7 ± 14.4	—	47.8 ± 11.7	—	—	48.9 ± 13.4	—	50.1 ± 13.2	—	—	47.4 ± 10.0	‖	51.5 ± 15.7	—	—
Impulsivity-hyperactivity	50.5 ± 12.2	—	48.2 ± 10.2	—	—	49.0 ± 11.4	—	50.1 ± 11.3	—	—	47.8 ± 9.8	—	51.2 ± 12.5	—	—
Verbal learning (CVLT)															
List A total recall	47.4 ± 11.2	‡	50.9 ± 14.0	‡	‖	50.2 ± 11.4	‡	48.0 ± 14.4	‡	—	48.0 ± 13.7	‡	50.6 ± 11.8	‡	—
Learning slope	−0.1 ± 1.0	—	−0.2 ± 1.2	—	—	−0.1 ± 1.0	—	−0.2 ± 1.1	—	—	−0.2 ± 1.2	—	−0.1 ± 1.0	—	—
Short delay free recall	−0.0 ± 1.0	—	−0.1 ± 1.3	—	—	0.0 ± 1.1	—	−0.2 ± 1.3	—	—	−0.2 ± 1.3	—	0.0 ± 1.0	—	—
Long delay free recall	−0.3 ± 1.2	—	0.0 ± 1.2	—	—	−0.1 ± 1.2	—	−0.2 ± 1.3	—	—	−0.2 ± 1.3	—	0.0 ± 1.1	—	—
Academics (WIAT)															
Math	99.5 ± 14.5	—	101.9 ± 16.5	—	—	102.8 ± 14.9	‖	97.7 ± 16.0	—	‖	101.9 ± 15.8	—	99.3 ± 15.2	—	—
Reading	102.5 ± 15.3	—	100.1 ± 14.7	—	—	105.4 ± 13.8	‡	95.9 ± 14.9	‖	‡	100.1 ± 15.5	—	102.7 ± 14.5	—	—
Spelling	103.2 ± 16.8	—	100.1 ± 15.9	—	—	105.6 ± 16.5	‡	96.6 ± 15.0	—	‡	99.4 ± 16.4	—	104.1 ± 16.2	—	—
Wechsler scales															
FSIQ	97.3 ± 17.8	—	100.6 ± 15.5	—	—	101.3 ± 15.5	—	96.2 ± 17.7	‖	—	99.6 ± 16.9	—	98.3 ± 16.5	—	—
Working memory (FFD)	95.1 ± 16.2	‖	97.3 ± 16.8	—	—	98.1 ± 15.3	—	92.7 ± 17.8	‖	—	95.1 ± 16.4	‖	97.0 ± 16.5	—	—
Processing speed (PSI)	104.9 ± 15.4	‖	100.4 ± 15.5	‖	—	108.1 ± 14.2	‡	95.7 ± 14.5	‖	‡	97.4 ± 15.3	—	107.6 ± 14.3	‖	‡

Abbreviations: CPT, Continuous Performance Test; CPRS, Connors Parent Rating Scales; CVLT, California Verbal Learning Test; FFD, Freedom from Distractibility Index; FSIQ, Full Scale IQ; PSI, Processing Speed Index; WIAT, Wechsler Individual Achievement Test.
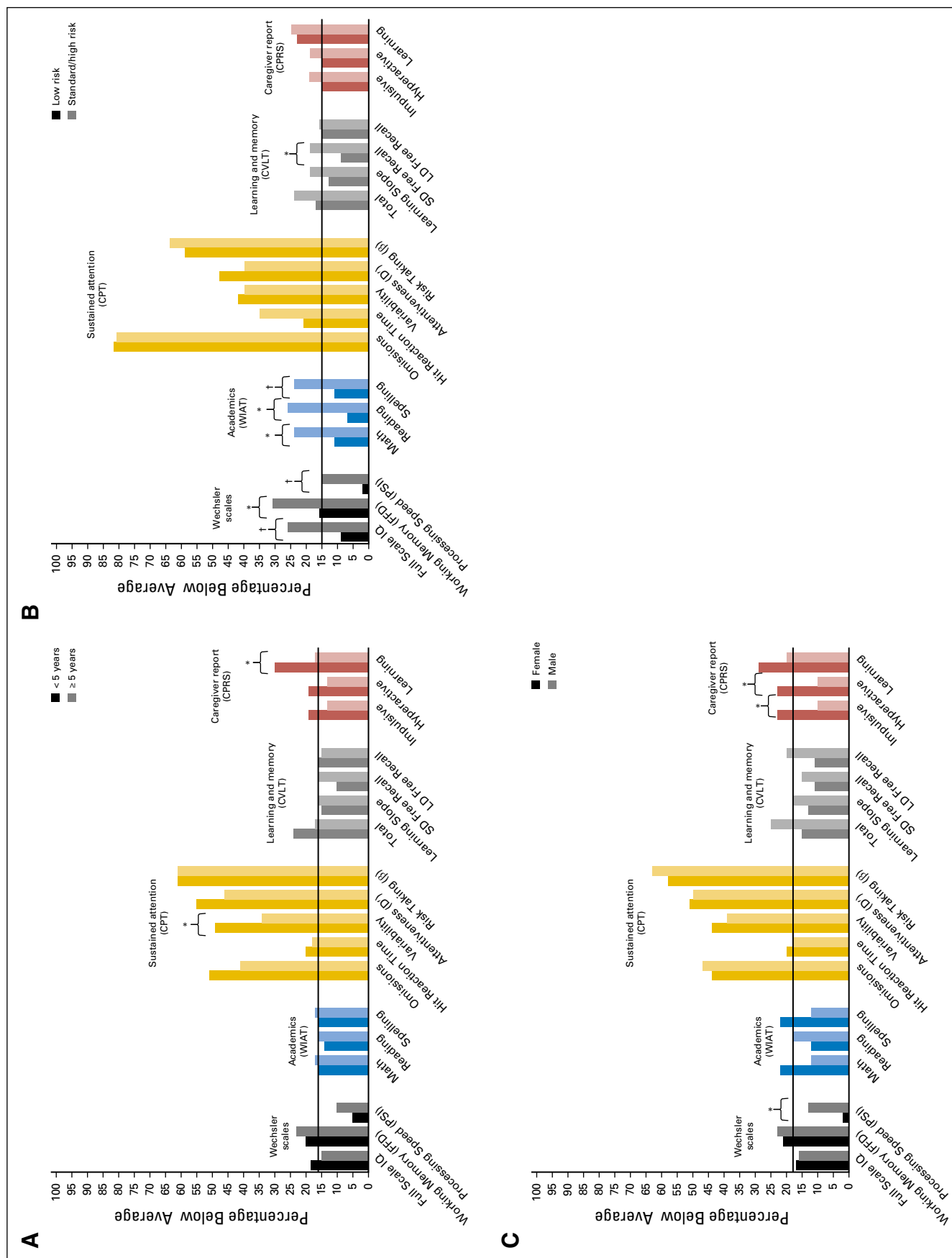*Two-sided P value from a one-sample t test comparing subgroup mean with normative values.
†Two-sided P value from independent samples t test comparing means between subgroups.
‡P ≤ .001.
§P ≤ .01.
‖P ≤ .05.

Table 3. Change in Neurocognitive Functioning From End of Therapy to 2 Years Post

	Intercept	P*	Slope	P†
Sustained attention (CPT)				
Attentiveness (D′)	60.12	‡	−1.45	
Hit reaction time (slow)	48.75		2.89	‡
Variability	58.94	‡	−0.51	
Risk-taking (β)	71.47	‡	−1.29	
Omissions	84.01	‡	−3.78	ǁ
Verbal learning (CVLT)				
List A total	49.19		0.22	
Learning slope	−0.24	§	0.10	
Short delay free recall	0.03		−0.10	
Long delay free recall	−0.11		0.03	
Wechsler scales				
Working memory (FFD)	94.86	‡	1.45	
Processing speed (PSI)	101.05		0.64	
Academics (WIAT)				
Math	99.43		1.06	
Reading	101.31		−0.81	
Spelling	99.32		1.98	ǁ

NOTE. Intercept is mean score for the overall group at the End of Therapy. Slope is estimated change in standardized score points from End of Therapy to 2 Years Post. Expected normative mean is 100 for estimated IQ, FFD, PSI, and WIAT; 50 for Connors Parent Rating Scales, CPT, and CVLT-List A Total; 0.0 for CVLT Learning Slope, Short Delay and Long Delay Free Recall.

Abbreviations: CPT, Continuous Performance Test; CVLT, California Verbal Learning Test; FFD, Freedom from Distractibility Index; PSI, Processing Speed Index; WIAT, Wechsler Individual Achievement Test.

*P value from one-sided t test comparing group mean to normative mean.

†P value from main effect of time.

‡Two-sided $P \leq .001$.

§Two-sided $P \leq .01$.

ǁTwo-sided $P \leq .05$.

standard/high risk: 65 μM), followed by standardized leucovorin rescue (low risk: 10 mg/m^2; standard/high risk: 15 mg/m^2), beginning 42 hours after the start of MTX infusion, every 6 hours for a total of five doses. Leucovorin dosage was increased if the MTX concentration was greater than 1.0 μM at 42 hours and continued until MTX concentration was less than 0.1 μM. This approach allowed us to clearly compare the effect of MTX dose between risk groups; however, the standardized rescue precluded an examination of the effect of leucovorin on neurocognitive outcome. Continuation treatment included weekly intravenous MTX at 40 mg/m^2 together with daily mercaptopurine for 3 weeks, followed by pulse therapy with vincristine plus dexamethasone at week 4 (low risk: 8 mg/m^2/d; high risk: 12 mg/m^2/d for 5 days). This treatment was 120 weeks for girls and 146 weeks for boys and was interrupted by two reinduction treatments. No patients received prophylactic CRT.

Neurocognitive Assessment

Assessments were conducted at week 6 of induction (Induction), 120 weeks post completion of consolidation (End of Therapy), and 2 years after completion of maintenance (2 Years Post), using standardized measures with demonstrated reliability and validity. Participants completed an age-appropriate measure of intelligence.[22-24] The Bayley Scales of Infant Development, 2nd edition[25] was administered to children ≤ 3.5 years old as a measure of global development. Indices of attention, working memory, and processing speed were also collected.[23,24] Patients ≥ 6 years

of age completed a computerized sustained attention measure,[26] which yields scores for omissions, reaction time, variability, vigilance, and risk taking. Caregivers completed standardized ratings of the impact of attention in daily life (eg, learning, hyperactivity, impulsivity).[27] Learning and memory and academic skills (reading, spelling, math) were assessed in patients ≥ 6 years old.[28-30] Measures were administered by master's-level psychological examiners under the supervision of a licensed psychologist.

Statistical Analyses

Descriptive analyses were conducted to characterize the group and to compare participants with and without data to establish representativeness. The percentage of the sample performing below the average range, operationalized as a mean score discrepancy of 1 SD from the normative sample, was calculated for each cognitive measure. χ^2 analyses were performed to compare the frequency of participants with below average performance to normative expectations. One-sample t tests were conducted to compare groups on the basis of age at diagnosis (< 5 or ≥ 5 years old), treatment risk arm (low or standard/high), and sex to normative means; cognitive scores for these subgroups were compared directly using two-sample t tests. Univariate and multivariable logistic regression were used to estimate the effect of age at diagnosis, risk arm, and sex on the probability of below average performance. Mixed-effects linear models were used to examine the change in cognitive scores over time for the overall group and in separate models with age at diagnosis, risk arm, and sex included as explanatory variables. Linear regression was used to examine whether End of Therapy attention predicted academics at 2 Years Post. All tests of statistical significance were two-sided. Data were analyzed using SAS version 9.3.

RESULTS

Participant Characteristics

Three hundred thirty-nine of 408 patients enrolled on Total XV Therapy completed at least one neurocognitive assessment. Of these, 243 (72%) completed End of Therapy, 211 (62%) completed 2 Years Post, and 167 (49%) completed End of Therapy and 2 Years Post. We have previously shown that there are no significant differences on relevant clinical and demographic variables between the 339 patients completing at least one assessment and the 69 patients with no assessment.[10] There were no significant differences between the cohort completing End of Therapy and 2 Years Post and the cohort completing End of Therapy or 2 Years Post. See Table 1 for demographics.

Neurocognitive Outcomes 2 Years Post

At 2 Years Post, the overall group did not significantly differ from normative expectations on measures of global intelligence (estimated IQ), academic skills (reading, math, or spelling), and learning and memory. There was a significantly higher frequency of below average performance compared with normative expectations (16%) on measures of sustained attention. The frequency of below average performance was 50% for attentiveness, 45% for omissions, and 41% for variability (P values < .001). Although mean

Fig 1. (A) Univariate logistic regression comparing the frequency of below average performance by age at diagnosis (< 5 years old; ≥ 5 years old). Reference group: ≥ 5 years old at diagnosis. (B) Univariate logistic regression comparing the frequency of below average performance by treatment risk arm (low; standard/high). Reference group: low risk. (C) Frequency of below average performance by sex. Univariate logistic regression comparing the frequency of below average performance by sex. Reference group: male. Black bars indicate the expected frequency of performance outside the average range in the normative sample (16th percentile). *$P \leq .05$, †$P \leq .01$, ‡$P \leq .001$. CPRS, Connors Parent Rating Scales; CPT, Continuous Performance Test; CVLT, California Verbal Learning Test; D′, discriminability; FFD, Freedom from Distractibility Index; LD, long delay; PSI, Processing Speed Index; SD, short delay; WIAT, Wechsler Individual Achievement Test.

scores were within normative expectations, caregivers reported a significantly greater frequency of learning problems in the overall group (P = .002). Results from bivariate correlation analyses showed a low association between performance and caregiver ratings of attention difficulties (Appendix Table A1, online only).

Table 2 shows results of mean comparisons examining the impact of age at diagnosis, treatment intensity, and sex on neurocognitive outcomes at 2 Years Post. Group means were within normative expectations on the majority of measures, with the exception of several indices of sustained attention. Children who were younger at diagnosis performed significantly worse than the older group on measures of attention variability and learning. Children treated with standard/high-risk therapy performed worse on measures of academics and processing speed compared with those treated with low-risk therapy. Processing speed was significantly lower in males than females.

Results of univariate logistic regression analyses are depicted in Figure 1. Younger age at diagnosis conferred increased risk for below

average performance on measures of attention variability (odds ratio [OR], 1.9; 95% CI, 1.0 to 3.4; P = .036) and increased caregiver-reported learning problems (OR, 2.1; 95% CI, 1.0 to 4.2; P = .049). Standard/high-risk treatment conferred increased risk for below average performance on measures of intelligence (OR, 0.3; 95% CI, 0.1 to 0.6; P = .002), working memory (OR, 0.4; 95% CI, 0.2 to 0.9; P = .035), processing speed (OR, 0.1; 95% CI, 0.0 to 0.6; P = .008), learning (OR, 0.4; 95% CI, 0.2 to 1.0; P = .047), and academics (math OR, 0.4; 95% CI, 0.2 to 0.8; P = .017; reading OR, 0.2; 95% CI, 0.1 to 0.6; P = .001; spelling OR, 0.4; 95% CI, 0.2 to 0.8; P = .017) compared with low-risk treatment. When compared with males, females had elevated caregiver ratings of hyperactivity (OR, 2.8; 95% CI, 1.2 to 6.4; P = .018) and impulsivity (OR, 3.0; 95% CI, 1.2 to 7.6; P = .021) and had less risk for slowed processing speed (OR, 0.2; 95% CI, 0.0 to 0.7; P = .018). Results of multivariable logistic regression showed females had elevated caregiver ratings of impulsivity (OR, 3.2; 95% CI, 1.2 to 8.6) and hyperactivity (OR, 2.9; 95% CI, 1.2 to 7.1; Appendix Table A2, online only).
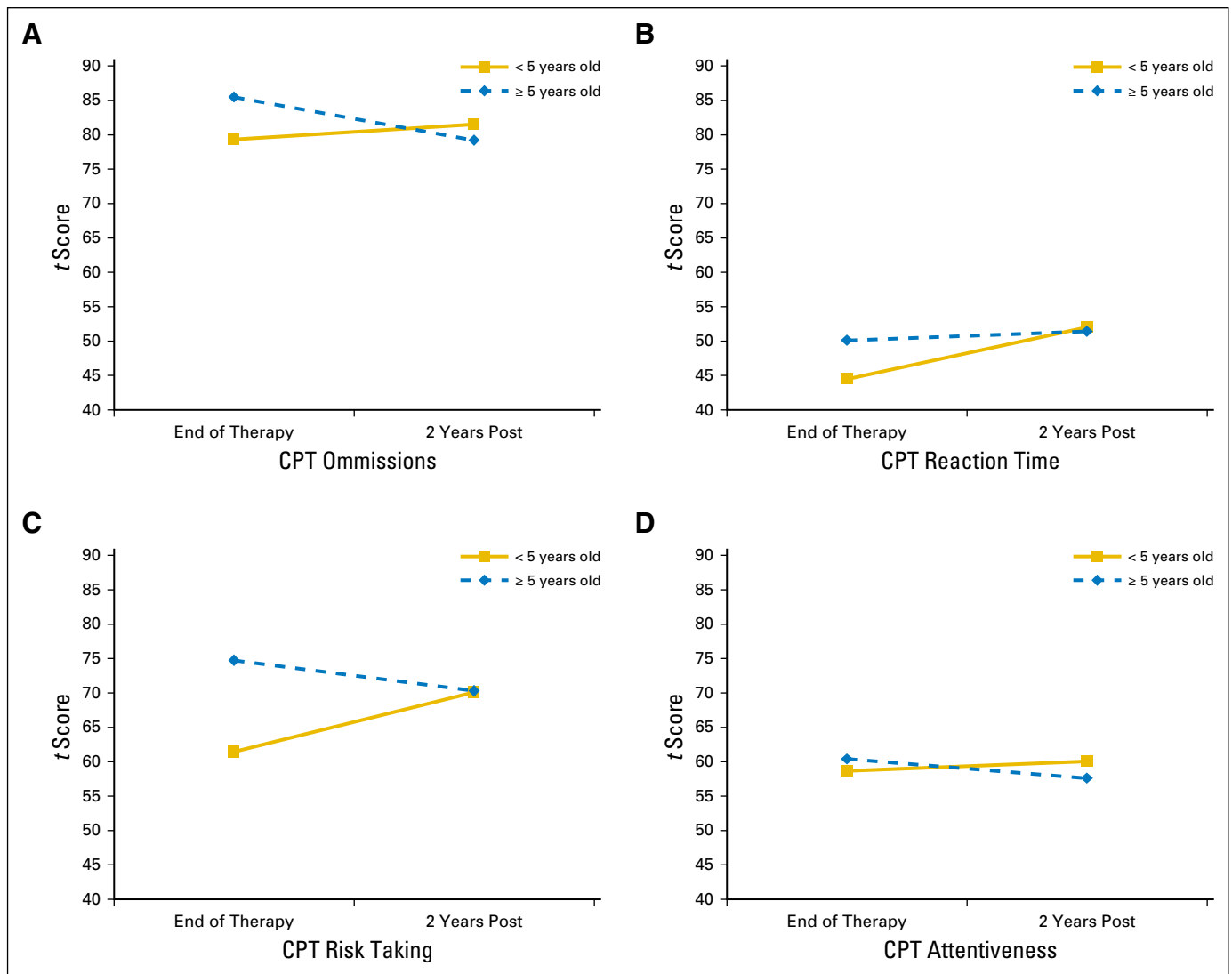
Fig 2. Change in sustained attention from End of Therapy to 2 Years Post by age at diagnosis. Group means from models with age at diagnosis predicting sustained attention scores (CPT): (A) omissions, (B) reaction time, (C) risk taking, and (D) attentiveness. Higher scores denote worse performance. CPT, Continuous Performance Test.

Change in Neurocognitive Functioning Over Time

We examined change in intelligence and caregiver ratings of attention from Induction to 2 Years Post. Although mean scores were within the average range at Induction and 2 Years Post, there was significant improvement in intelligence ($P = .040$) and significant increase in learning problems from Induction to 2 Years Post ($P = .006$). For all other variables, we examined change from End of Therapy to 2 Years Post (Table 3). The overall group continued to demonstrate sustained attention impairment at 2 Years Post, although performance slightly improved from End of Therapy to 2 Years Post. Significant improvement was evident in spelling scores. There were no significant changes over time on measures of working memory, processing speed, or memory; these are not considered in remaining subgroup analyses.

There were significant group differences in the rate of change by treatment intensity. Compared with patients treated with low-risk therapy, those treated on the standard/high-risk arm had a significantly greater rate of increase in caregiver-reported learning problems ($P = .034$). There were significant differences in the rate and direction of change by age at diagnosis and sex (Figs 2 and 3). In patients who were older at diagnosis, reaction time remained stable ($P = .007$), and omissions improved slightly ($P = .046$). In contrast, patients diagnosed at younger ages had stable performance in attentiveness ($P = .044$) and significantly increased problems over time in risk taking ($P = .002$). On caregiver-reported impulsivity and hyperactivity, females were rated as having significantly increased problems and males as having decreased problems over time ($P = .025$ and $.032$, respectively).

Predicting Intelligence and Academics

There were significant small- to moderate-sized associations between End of Therapy attention and 2 Year Post academics. At End of Therapy, slower reaction time significantly predicted improved academics at 2 Years Post (reading r, 0.34; 95% CI, 0.16 to 0.51; math r, 0.28; 95% CI, 0.09 to 0.45; and spelling r, 0.39; 95% CI, 0.21 to 0.55). At the End of Therapy, high variability

significantly predicted decreased reading (r, -0.23; 95% CI, 0.03 to 0.41) and estimated IQ (r, -0.25; 95% CI, 0.07 to 0.42) at 2 Years Post. End of Therapy caregiver ratings of learning and hyperactivity predicted lower reading (r, -0.20; 95% CI, 0.04 to 0.35; r, -0.17; 95% CI, 0.01 to 0.32) and math (r, -0.20; 95% CI, 0.04 to 0.35; r, -0.17; 95% CI, 0.01 to 0.33) scores at 2 Years Post.

DISCUSSION

We conducted a prospective, longitudinal study of neurocognitive outcomes in a representative cohort of childhood ALL survivors treated with contemporary therapy without prophylactic cranial irradiation. This study makes a significant contribution to the existing literature by addressing the limitations of previous work, including small sample sizes, treatment heterogeneity, and lack of consistency with regard to follow-up time points. Increased specification of the neurocognitive trajectory in the early survivorship phase will inform cognitive interventions designed to ameliorate difficulties and improve quality of life for the growing population of survivors.

Our findings support a priori hypotheses regarding improved outcomes for survivors treated with contemporary therapy. As a group, survivors performed within age expectations on several neurocognitive measures 2 years after completing therapy, confirming that omission of prophylactic CRT from childhood ALL therapy results in improved cognitive outcomes. Despite this notable improvement, our data show that survivors continue to demonstrate elevated risk for attention deficits. These difficulties are evident on performance measures and caregiver ratings, suggesting that these isolated deficits significantly and negatively impact real-world functioning. The low association between performance and caregiver ratings seen in our study is well-documented and underscores the importance of multi-modal measurement in order to gain the most comprehensive information, particularly given that both were found to predict academics.
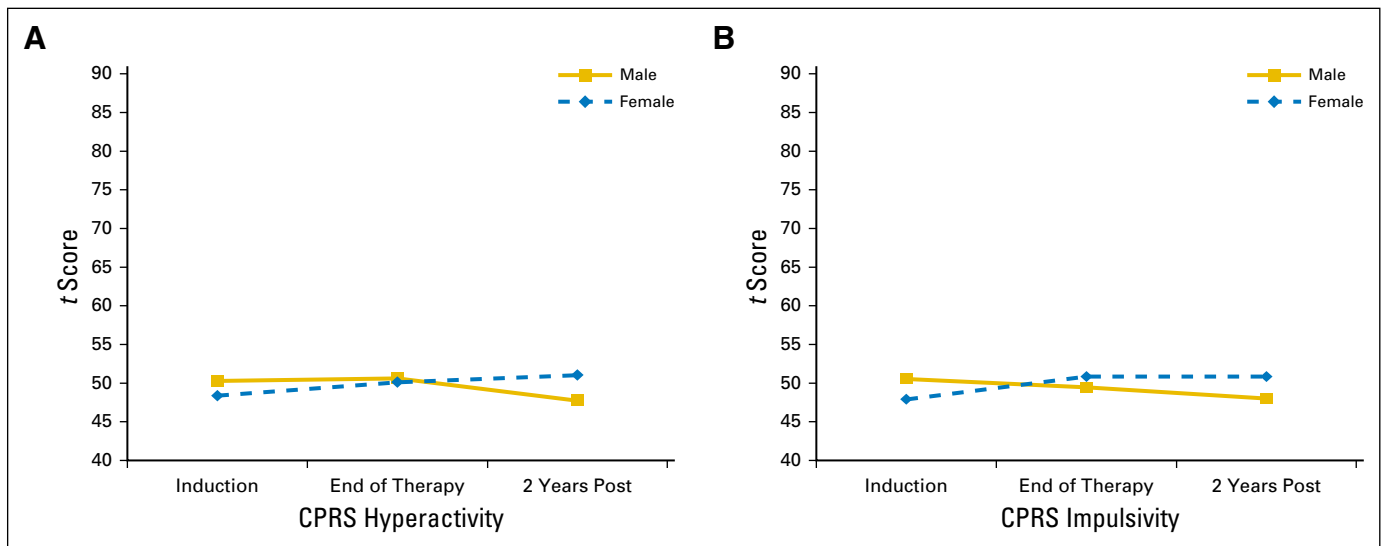
Fig 3. Change in caregiver ratings from End of Therapy to 2 Years Post by sex. Group means from models with sex predicting caregiver-reported problems (CPRS): (A) hyperactivity and (B) impulsivity. Higher scores denote more problems. CPRS, Connors Parent Rating Scales.

Our results suggest that treatment with higher-intensity CNS-directed chemotherapy and younger age at diagnosis continue to confer increased risk for neurocognitive difficulties 2 years after therapy. Structural neuroimaging studies of survivors show that treatment-related reductions in the volume and integrity of cerebral white matter predict worse performance on measures of attention and executive function, cognitive functions that are largely supported by the frontal cortex.[17,31] The development and maturation of the frontal lobe and white matter connectivity continue into early adulthood. As such, neurocognitive abilities supported by these mechanisms may be especially vulnerable to early insult.

The overall group of survivors demonstrated increased caregiver-reported learning problems from Induction to 2 Years Post; however, the rate of change was significantly greater for those treated with higher-intensity therapy. Age at diagnosis and sex significantly impacted the rate and direction of change. Problems with attention, impulsivity, and hyperactivity remain stable or increase for females and for those treated at younger ages. In contrast, inattention, hyperactivity, and impulsivity remained stable or decreased slightly for males. These findings are clinically informative, as they provide information regarding subgroups of patients at greatest risk. However, our findings of significantly increased risk for attention and learning problems in the overall group emphasize the importance of routine neurocognitive monitoring of all survivors treated with contemporary therapy. This is especially important in light of our finding that end-of-therapy attention problems predict subsequent academic difficulties. Although we found a small but significant improvement in attention performance over time, the overall group means remained substantially elevated at 2 Years Post, raising the possibility of regression to the mean.

Our findings regarding group differences by sex were somewhat unexpected. After accounting for the effect of age at diagnosis and treatment intensity, females continued to be at greater risk for impulsivity and hyperactivity. Longitudinal analysis revealed a pattern of increased difficulties in females and decreased difficulties in males. These findings are consistent with existing research that identifies female sex as a risk factor; however, we did not identify an appreciable impact of sex on neurocognitive ability at the end of therapy. Females may be at increased vulnerability because of sex-based differences in white matter development, including greater increase and later peak volume in healthy males.[32]

Our findings must be considered in the context of study limitations. Our longitudinal models included only survivors with data at each time point, raising the possibility of sample bias; however, we have demonstrated that there is no significant difference between patients completing one assessment and those completing no assessments, or between the cohort of patients completing assessments at End of Therapy and 2 Years Post and those completing End of Therapy or 2 Years Post. These data strongly support sample representativeness. Our study uses gold standard assessment measures with large and representative normative samples; however, inclusion of a medical control group would allow for isolation of effects of CNS-directed treatment from other illness-related experiences. Different versions of intelligence tests were needed, given the age range of our participants; although appropriate, this may have limited the sensitivity to detect change in intelligence over time. Findings regarding sex differences may be influenced by differences in caregiver expectations of behavior; however, this should have minimal impact on results, given that caregiver ratings are age and sex standardized. Given the large number of potential outcomes, we took care to restrict the potential for type 1 error by restricting our statistical comparisons to those outcomes that have been previously shown to be impacted in this cohort[10] and by restricting the longitudinal subgroup analyses to those domains where significant differences were found at the overall group level.

Future studies should consider including additional measures of everyday function (eg, teacher ratings or school grades) and measures of vulnerable domains (eg, attention) that are appropriate for very young children. Further specification of underlying neural or physiologic mechanisms of neurocognitive late effects is needed to promote early detection and to inform interventions to remediate early attention and behavior problems. Although shown to be efficacious, existing pharmacologic interventions have reduced acceptability among survivors of childhood cancer.[33] Given the young age at diagnosis in ALL, nonpharmacologic interventions should be designed and implemented from a developmental perspective and include components directed at survivors and family members.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at www.jco.org.

AUTHOR CONTRIBUTIONS

Conception and design: Kevin R. Krull, Ching-Hon Pui, Heather M. Conklin
Provision of study materials or patients: Ching-Hon Pui
Collection and assembly of data: Ching-Hon Pui
Data analysis and interpretation: Lisa M. Jacola, Kevin R. Krull, Deqing Pei, Cheng Cheng, Wilburn E. Reddick, Heather M. Conklin
Manuscript writing: All authors
Final approval of manuscript: All authors

REFERENCES

1. Kaatsch P: Epidemiology of childhood cancer. Cancer Treat Rev 36:277-285, 2010

2. Pui CH, Evans WE: A 50-year journey to cure childhood acute lymphoblastic leukemia. Semin Hematol 50:185-196, 2013

3. Jankovic M, Brouwers P, Valsecchi MG, et al; International Study Group on Psychosocial Aspects of Childhood Cancer: Association of 1800 cGy cranial irradiation with intellectual function in children with acute lymphoblastic leukaemia. ISPACC. Lancet 344: 224-227, 1994

4. Langer T, Martus P, Ottensmeier H, et al: CNS late-effects after ALL therapy in childhood. Part III: neuropsychological performance in long-term survivors of childhood ALL: Impairments of concentration, attention, and memory. Med Pediatr Oncol 38:320-328, 2002

5. Cheung YT, Krull KR: Neurocognitive outcomes in long-term survivors of childhood acute lymphoblastic leukemia treated on contemporary treatment protocols: A systematic review. Neurosci Biobehav Rev 53:108-120, 2015

6. Iyer NS, Balsamo LM, Bracken MB, et al: Chemotherapy-only treatment effects on long-term neurocognitive functioning in childhood ALL survivors: A review and meta-analysis. Blood 126: 346-353, 2015

7. Buizer AI, de Sonneville LM, van den Heuvel-Eibrink MM, et al: Behavioral and educational limitations after chemotherapy for childhood acute lymphoblastic leukemia or Wilms tumor. Cancer 106: 2067-2075, 2006

8. Kunin-Batson A, Kadan-Lottick N, Neglia JP: The contribution of neurocognitive functioning to quality of life after childhood acute lymphoblastic leukemia. Psychooncology 23:692-699, 2014

9. Buizer AI, de Sonneville LM, van den Heuvel-Eibrink MM, et al: Chemotherapy and attentional dysfunction in survivors of childhood acute lymphoblastic leukemia: Effect of treatment intensity. Pediatr Blood Cancer 45:281-290, 2005

10. Conklin HM, Krull KR, Reddick WE, et al: Cognitive outcomes following contemporary treatment without cranial irradiation for childhood acute lymphoblastic leukemia. J Natl Cancer Inst 104: 1386-1395, 2012

11. Buizer AI, De Sonneville LM, van den Heuvel-Eibrink MM, et al: Visuomotor control in survivors of childhood acute lymphoblastic leukemia treated with chemotherapy only. J Int Neuropsychol Soc 11: 554-565, 2005

12. Kingma A, van Dommelen RI, Mooyaart EL, et al: Slight cognitive impairment and magnetic resonance imaging abnormalities but normal school levels in children treated for acute lymphoblastic leukemia with chemotherapy only. J Pediatr 139: 413-420, 2001

13. von der Weid N, Mosimann I, Hirt A, et al: Intellectual outcome in children and adolescents with acute lymphoblastic leukaemia treated with chemotherapy alone: Age- and sex-related differences. Eur J Cancer 39:359-365, 2003

14. Jain N, Brouwers P, Okcu MF, et al: Sex-specific attention problems in long-term survivors of pediatric acute lymphoblastic leukemia. Cancer 115:4238-4245, 2009

15. Ashford J, Schoffstall C, Reddick WE, et al: Attention and working memory abilities in children treated for acute lymphoblastic leukemia. Cancer 116:4638-4645, 2010

16. Edelmann MN, Krull KR, Liu W, et al: Diffusion tensor imaging and neurocognition in survivors of childhood acute lymphoblastic leukaemia. Brain 137: 2973-2983, 2014

17. Reddick WE, Taghipour DJ, Glass JO, et al: Prognostic factors that increase the risk for reduced white matter volumes and deficits in attention and learning for survivors of childhood cancers. Pediatr Blood Cancer 61:1074-1079, 2014

18. Harila MJ, Winqvist S, Lanning M, et al: Progressive neurocognitive impairment in young adult survivors of childhood acute lymphoblastic leukemia. Pediatr Blood Cancer 53:156-161, 2009

19. Halsey C, Buck G, Richards S, et al: The impact of therapy for childhood acute lymphoblastic leukaemia on intelligence quotients: Results of the risk-stratified randomized central nervous system treatment trial MRC UKALL XI. J Hematol Oncol 4:42, 2011

20. Jansen NC, Kingma A, Schuitema A, et al: Neuropsychological outcome in chemotherapy-only-treated children with acute lymphoblastic leukemia. J Clin Oncol 26:3025-3030, 2008

21. Pui CH, Campana D, Pei D, et al: Treating childhood acute lymphoblastic leukemia without cranial irradiation. N Engl J Med 360:2730-2741, 2009

22. Wechsler D. Wechsler Preschool and Primary Scale of Intelligence, Revised. Psychological Corporation, 1989

23. Wechsler D. Wechsler Intelligence Scale for Children (ed 3). San Antonio, TX, Psychological Corporation, 1991

24. Wechsler D. Wechsler Adult Intelligence Scale (ed 3). San Antonio, TX, Psychological Corporation, 1997

25. Bayley N. Bayley Scales of Infant Development (ed 2). San Antonio, TX, Psychological Corporation, 1993

26. Conners CK: Connors' Continuous Performance Test. Toronto, Ontario, Multi-Health Systems, 2000

27. Conners CK: Connors' Rating Scales - Revised Technical Manual. North Tonawanda, NY, Multi-Health Systems, 2000

28. Delis DC, Kramer JH, Kaplan E, et al: California Verbal Learning Test, Adult Version. New York, NY, Psychological Corporation, 1987

29. Delis DC, Kramer JH, Kaplan E, et al: California Verbal Learning Test: Children's Version. New York, NY, Harcourt, 1994

30. Wechsler D. Wechsler Individual Achievement Test. New York, NY, Psychological Corporation, 1992

31. Reddick WE, Shan ZY, Glass JO, et al: Smaller white-matter volumes are associated with larger deficits in attention and learning among long-term survivors of acute lymphoblastic leukemia. Cancer 106:941-949, 2006

32. Schmithorst VJ, Holland SK, Dardzinski BJ: Developmental differences in white matter architecture between boys and girls. Hum Brain Mapp 29: 696-710, 2008

33. Conklin HM, Helton S, Ashford J, et al: Predicting methylphenidate response in long-term survivors of childhood cancer: A randomized, double-blind, placebo-controlled, crossover trial. J Pediatr Psychol 35:144-155, 2010

- - -

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Longitudinal Assessment of Neurocognitive Outcomes in Survivors of Childhood Acute Lymphoblastic Leukemia Treated on a Contemporary Chemotherapy Protocol

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or jco.ascopubs.org/site/ifc.

Lisa M. Jacola
No relationship to disclose

Kevin R. Krull
No relationship to disclose

Ching-Hon Pui
No relationship to disclose

Deqing Pei
No relationship to disclose

Cheng Cheng
Research Funding: Sigma Tau Pharmaceuticals (Inst)

Wilburn E. Reddick
No relationship to disclose

Heather M. Conklin
No relationship to disclose

Appendix

Table A1. Correlations Between Performance-Based and Caregiver-Reported Measures of Attention						
Measure of Attention					95% CI	
Caregiver Reported (CPRS)	Performance Based (CPT)	r	P		Lower	Upper
Learning	Omissions	0.13	.095		−0.02	0.28
	Discriminability (attention)	0.11	.169		−0.05	0.26
	Hit reaction time	−0.13	.098		−0.28	0.02
	Variability standard error	0.07	.347		−0.08	0.22
	β (risk taking)	0.11	.160		−0.04	0.26
Hyperactivity	Omissions	0.15	.054		−0.00	0.30
	Discriminability	0.10	.187		−0.05	0.25
	Hit reaction time	−0.12	.124		−0.27	0.03
	Variability standard error	0.08	.285		−0.07	0.23
	β	0.04	.620		−0.11	0.19
Impulsivity-hyperactivity	Omissions	0.15	.062		−0.01	0.29
	Discriminability	0.09	.262		−0.07	0.24
	Hit reaction time	−0.03	.670		−0.19	0.12
	Variability standard error	0.06	.476		−0.10	0.21
	β	0.02	.770		−0.13	0.18

Abbreviations: CPRS, Connors Parent Rating Scales; CPT, Continuous Performance Test.

Table A2. Multivariable Regression															
	Age at Diagnosis			Sex			Dexamethasone			ITMHA			HDMTX		
	< 5 Years v ≥ 5 Years*			Female v Male*			Per 100 mg/m^2			Per 50 mL			Per 5 g/m^2		
	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P
Attention (CPT)															
Hit reaction time	1.2	0.6 to 2.5	—	1.3	0.6 to 2.7	—	1.0	0.9 to 1.2	—	1.0	0.7 to 1.3	—	1.3	0.9 to 2.1	—
Variability	1.8	1.0 to 3.34	—	1.3	0.7 to 2.5	—	1.2	1.0 to 1.3	—	0.8	0.6 to 1.1	—	1.2	0.8 to 1.8	—
Omissions															
Attentiveness (D′)	1.5	0.8 to 2.7	—	1.0	0.6 to 1.9	—	1.0	0.9 to 1.1	—	0.8	0.6 to 1.1	—	1.2	0.8 to 1.8	—
β	1.2	0.6 to 2.1	—	0.9	0.5 to 1.6	—	1.0	0.9 to 1.1	—	0.9	0.7 to 1.2	—	1.2	0.8 to 1.7	—
Caregiver report (CPRS)															
Impulsive-hyperactive	1.7	0.7 to 4.3	—	3.2	1.2 to 8.6	†	1.1	0.9 to 1.3	—	1.0	0.7 to 1.4	—	1.3	0.7 to 2.2	—
Hyperactive	1.4	0.6 to 3.2	—	2.9	1.2 to 7.1	†	1.0	0.9 to 1.2	—	1.0	0.8 to 1.3	—	1.3	0.8 to 2.1	—
Learning	2.1	1.0 to 4.3	—	1.6	0.8 to 3.2	—	0.9	0.8 to 1.1	—	0.9	0.6 to 1.3	—	1.4	0.9 to 2.2	—
Verbal learning (CVLT)															
List A total	1.9	0.9 to 4.2	—	0.5	0.2 to 1.1	—	0.9	0.8 to 1.1	—	1.5	1.0 to 2.2	—	1.0	0.6 to 1.6	—
Learning slope	1.0	0.4 to 2.3	—	0.8	0.4 to 2.0	—	1.1	1.0 to 1.3	—	1.1	0.8 to 1.4	—	1.0	0.6 to 1.8	—
Short delay free recall	0.7	0.3 to 1.8	—	0.7	0.3 to 1.7	—	0.9	0.8 to 1.0	—	1.0	0.7 to 1.4	—	1.4	0.8 to 2.4	—
Long delay free recall	1.3	0.6 to 3.1	—	0.4	0.2 to 1.0	—	0.9	0.8 to 1.0	—	1.2	0.9 to 1.5	—	1.1	0.7 to 1.8	—
Academics (WIAT)															
Math	0.9	0.4 to 2.1	—	2.6	1.1 to 6.0	—	1.0	0.8 to 1.1	—	1.4	0.9 to 2.0	—	1.4	0.8 to 2.3	—
Reading	1.1	0.5 to 2.5	—	0.7	0.3 to 1.8	—	1.0	0.9 to 1.1	—	1.4	1.0 to 1.9	—	1.5	0.9 to 2.6	—
Spelling	0.9	0.4 to 2.1	—	2.6	1.1 to 6.0	—	1.0	0.8 to 1.1	—	1.4	0.9 to 2.0	—	1.4	0.8 to 2.3	—
Wechsler scales															
FSIQ	1.4	0.6 to 3.0	—	1.4	0.6 to 3.2	—	1.1	1.0 to 1.3	—	1.3	1.0 to 1.7	—	1.5	0.9 to 2.5	—
Working memory (FFD)	1.0	0.4 to 2.2	—	1.3	0.5 to 3.0	—	1.3	1.0 to 1.6	—	1.3	1.0 to 1.8	—	1.1	0.6 to 1.9	—
Processing speed (PSI)	0.8	0.2 to 2.8	—	0.2	0.0 to 0.8	—	0.9	0.7 to 1.1	—	1.1	0.6 to 1.8	—	2.4	1.1 to 5.3	—

NOTE. Two-sided P values are from multiple logistic regression models adjusted for all other depicted variables.
Abbreviations: CPRS, Connors Parent Rating Scales; CPT, Continuous Performance Test; CVLT, California Verbal Learning Test; FFD, Freedom from Distraction Index; FSIQ, Full Scale IQ; HDMTX, high-dose IV methotrexate; ITMHA, intrathecal methotrexate, cytarabine, hydrocortisone; OR, odds ratio; PSI, Processing Speed Index; WIAT, Wechsler Individual Academic Tests.
*Referent group for statistical comparison.
†$P ≤ .05$.