Published in final edited form as:

J Pediatr Hematol Oncol. 2014 January ; 36(1): 8–15. doi:10.1097/MPH.0000000000000000.

Neurocognitive and Neuroradiologic Central Nervous System Late Effects in Children Treated on Pediatric Oncology Group (POG) P9605 (standard risk) and P9201 (lesser risk) Acute Lymphoblastic Leukemia Protocols (ACCL0131): A Methotrexate Consequence? A Report from the Children's Oncology Group

Patricia K Duffner, MD[1], **F Daniel Armstrong, PhD**[2], **Lu Chen, PhD**[3], **Kathleen Helton, MD**[4], **Martin L Brecher, MD**[5], **Beverly Bell, MD**[6], and **Allen R Chauvenet, MD, PhD**[7]

[1]Department of Neurology, University at Buffalo School of Medicine, Buffalo, New York

[2]Department of Pediatrics, University of Miami Miller School of Medicine and Holtz Children's Hospital, Miami Florida

[3]Statistics and Data Center, Children's Oncology Group, Monrovia California

[4]Department of Radiological Sciences, St Jude Children's Research Hospital, Memphis Tennessee

[5]Department of Pediatrics, Roswell Park Cancer Institute, Buffalo, New York

[6]Department of Pediatrics, Georgia Health Science Center, Augusta Georgia

[7]Department of Pediatrics, West Virginia University Health Science Center, Charleston, West Virginia

Abstract

Concerns about long-term methotrexate (MTX) neurotoxicity in the 1990s led to modifications in intrathecal therapy, leukovorin rescue and frequency of systemic MTX administration in children with acute lymphoblastic leukemia. In this study, neurocognitive outcomes and neuroradiologic evidence of leukoencephalopathy were compared in children treated with intense central nervous system directed therapy (P9605) versus those receiving fewer CNS-directed treatment days during intensive consolidation (P9201). 66 children from 16 Pediatric Oncology Group (POG) institutions with “standard risk” ALL, 1.00 to 9.99 years at diagnosis, without evidence of CNS leukemia at diagnosis were enrolled on ACCL0131: 28 from P9201 and 38 from P9605. MRI scans and standard neuropsychologic tests were performed > 2.6 years following end of treatment. Significantly more P9605 patients developed leukoencephalopathy than P9201 (68%, 95% CI 49%-83% vs. 22%, 95% CI 5%-44%; p=0.001) identified as late as 7.7 years following end of treatment. Overall 40% of patients scored <85 on either VIQ or PIQ. Children on both studies had significant attention problems but P9605 children scored below average on more neurocognitive

Corresponding Author Patricia K Duffner, MD, 39 Saybrook Place, Buffalo, New York 14209, Phone 716-884-9484, FAX-716-849-6651, duffner@buffalo.edu.

Conflicts of interest and Source of funding: None are declared. Data from this paper have not been previously presented or published.

measures than those treated on P9201 (82%, 14/17 measures vs. 24%, 4/17 measures). This supports ongoing concerns about intensive methotrexate exposure as a major contributor to CNS late effects.

Keywords

methotrexate leukoencephalopathy; cognitive late effects

Introduction

B-precursor acute lymphoblastic leukemia (ALL) is the most common childhood malignancy in industrialized countries, with approximately 3-4 children/100,000 diagnosed annually. (1) Combination chemotherapy and aggressive central nervous system (CNS) prophylaxis have resulted in cure rates approaching 90% for patients with standard risk ALL. (2, 3) An important component of ALL therapy is the antimetabolite methotrexate (MTX). As intravenous doses of MTX have increased over the past two decades, concerns about acute and chronic effects of MTX therapies have grown, particularly those affecting the CNS and cognitive function. (4, 5, 6)

As early as 1978, changes on CT scans of children with ALL were attributed to MTX leukoencephalopathy. (7) Subsequent studies suggested that MTX alone, in the absence of CNS leukemia or cranial irradiation, was not a major risk factor for leukoencephalopathy, although cognitive impairment was reported as children completed therapy. (8, 9) As treatment intensity with MTX increased on POG study 9005, reports of seizures, dementia and changes on neuroimaging also increased. Two arms of the study included intravenous (IV) MTX 1 gm/m2 while one used oral MTX (0.3gms/m2). Acute neurotoxicity was significantly greater on the two study arms with IV MTX (8.3% and 11.2%) compared with the arm which used oral MTX (3.7%). Leukoencephalopathy was identified on either CT or MRI in 75% and 77.1% of symptomatic patients treated on the IV-MTX arms compared with 15% of patients in the oral MTX arm. Intensification using 12 courses of repeated IV MTX in the setting of low dose leukovorin rescue was identified as the likely risk factor. (10)

Prior to the availability of the P9005 results, P9405, the next POG study for standard risk ALL, utilized 12 courses of IV MTX in a regimen similar to 9005, but was closed due to unacceptable acute neurotoxicity. The next study, P9605, reduced the number of IV MTX courses from 12 to 6. This regimen was considered to be less neurotoxic as it paralleled the IV MTX courses given on the lesser-risk study P9201. In contrast to P9005, children treated on P9201, a treatment regimen for standard risk children with favorable cytogenetics, had little evidence of acute neurotoxicity (3%). (11)

In 2002, in response to concerns raised by investigators at the Roswell Park Cancer Institute, a study (ACCL0131) was mounted to compare neuroradiologic evidence of leukoencephalopathy and neurocognitive deficits in children treated on P9605 and P9201. Although (a) induction on P9605 and P9201 was the same, (b) both employed the same doses and timing of IV MTX and leucovorin during consolidation, and (c) no child received

radiation therapy, there were 3 differences in the treatment regimens. The first difference related to the use of TIT (triple intrathecal therapy) vs. ITM (intrathecal methotrexate). Patients on P9605 received TIT for induction, consolidation and maintenance until an amendment in July 1999 when all TIT was replaced with ITM (due to an excessive incidence of seizures). Patients on P9201 also received TIT during induction, consolidation and maintenance until an amendment in August 1996 when all post-induction TIT was switched to ITM. In July 1999, TIT was replaced with ITM during induction for patients on P9201 as well. The second difference is highlighted in Table 1. Patients on P9605 had 5 lumbar punctures with TIT (or ITM) delivered a week prior to IVMTX, (hence without leukovorin rescue), while those on P9201 had TIT (or ITM) concurrent with IVMTX, and were consequently protected by leukovorin rescue. This created a major difference in CNS exposure to MTX. In addition, patients on P9605 received MTX 2 out of every 3 weeks (either intravenously or intrathecally) on P9605 whereas patients on P9201 received no MTX for 2 weeks between doses. Finally, while patients on both P9201 and P9605 received induction, consolidation and maintenance chemotherapy, patients on P9605 also received post-consolidation intensive therapy. Children on regimen one of the study received standard 6 MP (mercaptopurine) (75 mg/m2 Q hs) + standard MTX (20 mg/m2 IM) weekly ; regimen 2 received standard 6 MP+ divided dose MTX (25 mg/m2 PO Q 6 hours X 4 every other week); regimen 3 received standard MTX with divided dose 6MP (37.5 mg/m2 po bid qd) and regimen 4 received divided dose MTX and divided dose 6MP. These protocols offer the unique opportunity to document whether the treatment differences impacted long-term radiologic, clinical and cognitive neurotoxicity.

Method and Materials

To be eligible for ACCL0131, patients had to be “good prognosis” by NCI risk, 1.00 to 9.99 years old at diagnosis, and without evidence of CNS3 at diagnosis. Patients had to be registered after the opening of P9605 (i.e. after 4-1-1996 but before amendment #8 dated 7-29-1999 or registered on P9201 after amendment #6 (8-26-1996). Patient assignment to P9201 was based on cytogenetic findings in the diagnostic bone marrow, i.e. the presence of Trisomy 4 and 10. ACCL0131 was later amended to exclude patients who had relapsed and those with Down syndrome and to include data and materials obtained a minimum of 3 years following completion of therapy.

The study was initially limited to selected POG institutions and patients that indicated prior to study initiation that the imaging and neuropsychological studies could be performed. From these, a randomly selected list of potentially eligible patients was identified by the statistics center; the selected centers/patients were invited to participate in the study. Patients from Roswell Park Cancer Institute were not initially invited to participate since this site had been the first to recognize that their patients on P9605 and P9201 differed in terms of neuroimaging and neuropsychological outcomes. Later, through an amendment, the study was opened to COG group-wide and to all eligible patients. When the study was opened group-wide, the Roswell patients were included.

Local IRB approval and informed consent were obtained on all patients. Mandated studies included MRI scans of the brain and a series of standardized neuropsychological tests with population norms. (appendix)

Magnetic resonance imaging

The MRI scans consisted of axial T1, T2 and fluid attenuated inversion recovery (FLAIR) images of entire brain. Coronal scans were not included due to cost considerations. No contrast was administered. Sedation was used based on the age of the patient and anticipated level of cooperation. All MRI scans were reviewed centrally and concurrently by a neuroradiologist (KJH) with experience in grading leukoencephalpathy , and a pediatric neurologist. (12) Both reviewers were blinded to the patients' treatment protocol. Although there was discussion about each case, there was no discordance. Scans were coded as normal/abnormal. Abnormal included the presence of white matter changes compatible with leukoencephalopathy.

Leukoencephalopathy was defined as abnormal T2 weighted hyperintensities in the deep white matter (centrum semiovale, periventricular white matter). Leukoencephalopathy was further coded by grade, as modified by Chu et al as mild (mild diffuse T2 hyperintensities in the periventricular white matter/centrum semiovale), moderate (moderate T2 hyperintensities which extend almost to the gray-white junction) or severe (severe T2 hyperintensities involving confluent deep white matter from the frontal horns to the trigone, which may extend to the subcortical U fibers). (13) Location was defined as anterior, posterior or both. In addition the presence or absence of hydrocephalus, calcification and focal lesions were noted.

Neurocognitive Testing

Neurocognitive tests were administered by a licensed psychologist or a trained intern, fellow, or psychometrician under the direct supervision of a licensed psychologist. The specific measures, area of function assessed, and test characteristics are included in the appendix.

Study Design and Statistical Analyses

The primary aim was to compare the prevalence of neurocognitive abnormalities in children treated on P9605 vs. P9201. The primary endpoint was the percentage of patients with intellectual deficit (either Verbal (VIQ) or Performance IQ (PIQ) below 85), with planning parameters of 20% VIQ or PIQ deficit for P9201 and 40% for P9605. The original design yielded 80% power with 1-sided alpha level of 0.05, with accrual in two-stages: Stage 1—30 and 60 patients from P9201 and P9605 respectively; Stage 2—24 and 48 patients from P9201 and P9605 respectively.

The study was closed due to poor accrual before meeting the targets for Stage 1. With the final number of patients for primary analysis of 16 and 31 respectively for P9201 and P9605, the power for detecting a 20% vs. 40% intelligence deficit would have been 39%. The original design used a one-sided statistical test since the difference in only one direction (superiority of P9201) was of interest; there was no interest in establishing the inferiority of

P9201 and no interest in differentiating between P9201 being comparable to P9605 and P9201 being worse than P9605.

Due to lack of accrual, all analyses performed were exploratory in nature. No multiple comparison adjustments were made in any analysis. All P values were based on two sided tests except the analysis comparing the frequency of intellectual deficit between P9201 and P9605 which was based on one-sided test per original design as indicated.

Two-sample t-tests were used to compare the mean of continuous variables (such as age at diagnosis or neurocognitive test score.) Comparisons of categorical variables (such as gender or presence/absence of leukoencephalopathy) between two groups were based on Fisher's exact test except that Chi-square test was used for comparison of race groups. Two-sided exact 95% confidence interval was computed for binomial proportions. The exact test of binomial proportions was used to compare the observed percentage of patients who scored greater than one standard deviation worse than the normative mean to the expected percentage in the general population, which is 16%.

Results

Participants

ACCL0131 was active from July 2002 to June 2007. The study closed due to slow accrual before reaching the stage 1 target of 30 patients from P9201 and 60 from P9605.

Patients (N=66) were enrolled from 16 institutions; 28 were enrolled on P9201 and 38 on P9605 between 1996 and 1999. One patient, treated on P9201, relapsed prior to enrollment on ACCL0131 and was excluded in all the analyses. A second patient with relapse was included in the analysis because the neuropsychological testing and MRI scan were performed prior to relapse. Four of 28 patients on P9201 received TIT rather than ITM due to statistical office/institutional error. Twenty-one of 38 patients on P9605 received both TIT and ITM, the amount of each depending on when they entered the treatment protocol. Six enrolled patients had neither MRI nor neuropsychological testing data available for review by the time the study was closed and are not included in any of the analyses. Thus there were 59 evaluable patients (24 from P9201 and 35 from P9605). (Table 2)

Under the original plan, 110 patients were selected from 15 POG institutions for Stage 1 (target n=90). Three of the sites never received IRB approval, eliminating 16 patients from the “invited “group. The study was later opened group-wide because of poor enrollment. Altogether 43 enrolled patients were initially invited; of those 7 were non-evaluable due to relapse or no MRI/ neuropsychological testing, leaving 36 evaluable patients. Another 23 enrolled and evaluable patients were not “invited” and were convenience samples.

MRI Results

Fifty-four patients had MRI scans performed; 23 from P9201 and 31 from P9605. There was a significant difference in the presence of leukoencephalopathy; P9201 with 5/23 (22%, 95% CI 5%-44%) vs. P9605 with 21/31 (68%, 95% CI 49%-83%) (p=0.001). With one exception, all leukoencephalopathy positive scans, (on both the T2 and FLAIR sequences),

were considered mild. One patient with moderate leukoencephalopathy on both the T2 and FLAIR sequences was treated on P9605. There were no cases of hydrocephalus identified. Calcification, which was considered indeterminate, was identified in one patient who had been treated on P9605. Focal abnormalities were identified on T1, T2 and FLAIR sequences. More focal lesions occurred on P9605 (11/31) vs. P9201 (2/23), p=0.03 on T2 and FLAIR sequences for which the result (lesion yes/no) was the same on all patients.

Neurocognitive Outcomes

Neurocognitive evaluations were completed for 52 patients. Not all patients completed all the subtests, thus the actual number of patients for different subtests varied. Scores more than one standard deviation from the normative mean (below 85 for standard and index scores, below 7 for scaled scores, above 60 for T-Scores) were considered clinically meaningful, since these represented below average to impaired performance. In the general population, 16% of children would be expected to score in this range.

For the primary study endpoint, there was no significant difference in the number of patients scoring <85 on the age appropriate Wechsler VIQ or PIQ (5/16 on P9201 (31%) vs. 14/31 on P9605 (45%), one sided p=0.27. Due to lack of accrual, the power to detect a statistically significant difference is compromised. A strong trend toward a significant mean Full Scale IQ (FSIQ) difference [P9605 (mean = 90.1) vs. P9201 (mean = 100.4]; (p= 0.06) was observed. Of concern, 40% (19/47) of patients in this study had either VIQ or PIQ scores below the normative range.

In exploratory analyses, no significant differences were identified between males and females in the percentage that scored below 85 on either VIQ or PIQ. However 42% of males (11/26) and 38% of females (8/21) scored more than 1 SD below the normative mean of the general population.

The two study cohorts were compared on their performance on specific measures of function. In these exploratory analyses, significant differences were found for the Wechsler Perceptual Organization Index (P=0.03), the Woodcock-Johnson Revised Processing Speed Cluster (P=0.05), and the Beery Test of Visuomotor Integration (P=0.03) with patients on P9605 having lower mean scores than those on P9201. An exception to this pattern was found for the Commission Errors of the Connors Continuous Performance Tasks (CPT), where patients treated on P9201 performed more poorly than those treated on P9605. (Table 3)

In a separate exploratory analysis, the percent of patients in each study who scored > 1 standard deviation (SD) worse than the normative mean were compared with the expected 16% of the normal population on the VIQ, PIQ and FSIQ as well as on 14 subtests (excluding the Connors CPT). More than 16% of patients treated on P9605 scored >1 SD worse than the normative mean on 14/17 (82%) specific measures of neurocognitive function compared to 4/17 (24%) specific measures for patients treated on P9201. (Table 4)

High T scores (>60) are indicative of clinically significant attention problems on the Connors CPT. In both studies more than 16% of patients scored > 1 SD worse than the

normative mean on the majority of CPT subtests (11/12 subtests on P9201 and 10/12 on P9605). (Table 4)

Comparison of MRI Leukoencephalopathy and IQ results

Both leukoencephalopathy results and IQ data were available for 43 participants. In exploratory analyses, there was no significant difference (p=0.36) in VIQ or PIQ <85 based on leukoencephalopathy (11/22 with leukoencephalopathy ; 7/21 without.) There were also no statistically significant differences for mean FSIQ, VIQ or PIQ based on leukoencephalopathy.

There were significant differences on means of specific measures of attention (commissions (P=0.05), 45.0 vs. 51.5, and omissions (P=0.04), 48.7 vs. 67.2) based on presence of leukoencephalopathy , with patients having leukoencephalopathy performing significantly worse on these subtests.

Among patients with neurocognitive scores >1 SD worse than the normative mean on a particular measure, the percentage with leukoencephalopathy was 80% for VIQ (8/10), 64% for PIQ (9/14) and 89% for FSIQ (8/9). For P9605, the percent with leukoencephalopathy among those scoring worse than the normative mean was 87.5% (7/8) for VIQ, 81.8% (9/11) for PIQ, and 100% (8/8) for FSIQ. For tests measuring verbal comprehension, perceptual organization, processing speed, freedom from distractibility, memory and visual-motor integration, between 75% and 100% of patients treated on P9605 with abnormal performance (> 1 SD worse than the normative mean) also had leukoencephalopathy.

Discussion

The goal of this study was to determine whether there were differences in neurocognitive function and neuroradiologic findings of leukoencephalopathy in children treated on P9605 and P9201.

We report a number of important findings. First, significantly more patients treated on P9605 had leukoencephalopathy than on P9201. Whereas some authors have questioned the significance of transient white matter changes, white matter changes in this study were not transient, with some identified as late as 7.7 years after completion of treatment. (14) A report of a small, single institution study of children undergoing active treatment on P9605 found that 78% of patients had white matter changes on at least one MRI. (15) When their short term results are viewed in the context of the long term outcomes of patients on P9605, persistent white matter changes are evident in the majority.

Second, while there were no statistically significant differences between the studies in the number of children with VIQ or PIQ < 85 (> 1 SD worse than the normative mean of the general population), 40% of the total group of patients treated on P9605 and P9201 had PIQ or VIQ scores that were abnormal compared to the general population. Neurocognitive late effects related to ALL treatment have been repeatedly identified in the specific areas of attention, freedom from distractibility, speed of information processing, verbal memory, visual memory, verbal comprehension, visuospatial skills, visual motor integration and

executive function. (16, 17, 18, 19, 20, 21, 22). We found similar areas of neurocognitive weakness with the exception of verbal memory and executive function.

Third, there was a strong relationship between MRI evidence of leukoencephalopathy and neurocognitive function. While there were no differences between those with/without leukoencephalopathy on measures of global cognitive function (FSIQ, VIQ, PIQ), the majority (75% - 100%) of children treated on P9605 who scored > 1 SD worse than the normative mean on areas of specific neurocognitive function also had leukoencephalopathy. The difficulty with interpreting this data is that most of the patients treated on P9605, with and without abnormal cognitive testing, had leukoencephalopathy. Whether larger numbers of patients in this cohort would have yielded statistically significant results is unknown. In addition, conventional MRI > 5 years from diagnosis may not adequately reflect lasting neurocognitive effects of acute white matter injury, not does it quantify changes due to myelin or axonal injury. Other imaging approaches, such as white matter anisotropy or diffusion, and structural /functional white matter connectivity should be considered to advance understanding in subsequent studies. (16)

In summary, patients treated on P9605 fared significantly worse in terms of leukoencephalopathy than patients treated on P9201. In addition, patients on P9605 performed worse than the general population (greater than 16% scored > 1 SD worse than the normative mean) on more tests of neuropsychological function than those treated on P9201, with the exception of the Connors CPT where both groups performed worse than the general population. This suggests that differences in protocol design of P9201 reduced the incidence of both leukoencephalopathy and neurocognitive damage compared to P9605. There were 3 important differences in the treatment regimens of the 2 protocols. The first involved the administration of TIT vs ITM during consolidation and maintenance. Although both protocols were amended in 1999 to replace all TIT with ITM , prior to that patients on P9605 received TIT throughout treatment whereas patients from P9201 who were eligible for ACCL0131 (i.e. after amendment August 1996) received TIT only during induction. This issue is complex however, since some patients were entered on treatment before the 1999 amendment and continued treatment afterwards—thus receiving both TIT and ITM during their course of treatment. It is unlikely however that the difference in IT medications influenced long term neurototoxicity since Kadan-Lottick and colleagues found no significant differences in neuropsychological function between children treated with TIT vs. IT MTX. (23)

It is likely that the differences in intensive CNS therapy during consolidation were a more important contributing factor, i.e. the administration of leucovorin following all IT courses for patients on P9201 (hence reducing CNS exposure) and the difference in the timing of MTX administration. Patients on P9605 received MTX, either IV or IT, 2 out of every 3 weeks while patients on P9201 had 2 weeks between courses. It is possible that the frequent administration of “unprotected” IT MTX in P9605 may have depleted folic acid stores in the central nervous system and not allowed the brain to recover sufficiently before the next dose was given.

. The significance of the third difference between the treatment regimes, i.e. the post-consolidation intensive therapy on P9605, is more difficult to interpret because of the multiplicity of regimens and the limited number of patients. The possible role of divided dose MTX and/or 6MP in 3 of the 4 arms needs clarification in future studies.

Finally, an area of particular weakness for patients treated on both studies was in attention and distractibility, deficits that are likely to have an impact on future educational success. The fact that these impairments were found in both groups suggests that there may be a subgroup of children with genetic risk factors (e.g. polymorphisms related to folate metabolism) that make them particularly vulnerable to the adverse effects of MTX. (24) Since CSF folate levels were not requested, this is speculative only.

While the results of this study are concerning, conclusions are limited by the relatively small number of patients. Unfortunately, this is a frequent limitation of late effects studies even when the cost of neuropsychologic studies is covered. (25) In addition it is possible that selection bias may have been a factor in the results. It is not clear why some families refused to enter their children in the study while others agreed to do so. It is possible, although by no means certain, that those who had children with academic issues may have been more interested in pursuing the neuropsychological testing.

In 2002, an editorial was written in the Journal of Pediatric Hematology /Oncology entitled “More good news about neuropsychological late effects in long-term survivors of acute lymphoblastic leukemia.” (26) While it is true that children with ALL treated with cranial irradiation for CNS prophylaxis during the 1980s had far greater neuropsychological deficits, our results suggest that the chemotherapy protocols of the late 1990s were far from benign. Current COG protocols do not give intensive/frequent IT MTX interspaced with high dose IV MTX, a practice our data would support.

Acknowledgments

Research is supported by the Chair's Grant CA98543-08 of the Children's Oncology Group from the National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NCI or the NIH.

A complete listing of grant support for research conducted by CCG and POG before initiation of the COG grant in 2003 is available online at: http://applications.childrensoncologygroup.org/admin/grantinfo.htm”

Appendix

Appendix

Neurocognitive areas of function, tests, and test characteristics

Functional Area	Test	Indices & Subtests	Normative M/SD
Global Intellectual Function	Wechsler Preschool & Primary Intelligence Scale, Revised (WPPSI-R) for ages 4-6 years	Full Scale IQ (FSIQ), Verbal IQ (VIQ), Performance IQ (PIQ)	M=100, SD=15[a]
	Wechsler Intelligence Scale for Children, 3rd Edition (WISC-III) for ages 6-16 years		

Functional Area	Test	Indices & Subtests	Normative M/SD
Verbal Abilities	WPPSI-R or WISC-III	Verbal Comprehension Index (VCI)	M=100, SD=15[a]
Perceptual and Spatial Abilities	WPPSI-R or WISC-III	Perceptual Organization Index (POI)	M=100, SD=15[a]
Spatial Planning	NEPSY	Tower Test Scaled Score	M=10, SD=3[b]
	Beery Test of Visual-Motor Integration (VMI)	VMI Standard Score	M=100, SD=15[a]
Attention & Concentration	WPPSI-R or WISC-III	Freedom from Distractibility Index (FDI)	M=100, SD=15[a]
	Conners Continuous Performance Test (CPT)	Omission Errors, Commission Errors, Inattentiveness, Impulsivity, Variability, Reaction Time, Risk, Vigilance Scales	M=50, SD=10[c]
Processing Speed	WPPSI-R or WISC-III	Processing Speed Index (PSI)	M=100, SD=15[a]
	Woodcock-Johnson, Revised (WJR)	Processing Speed Cluster	M=100, SD=15[a]
Memory	Wide Range Assessment of Memory and Learning (WRAML)	Overall Memory Index Memory Subtests: Verbal Learning, Picture Memory, Story Memory, Design Memory	M=100, SD=15[a] M=10, SD=3[b]

[a] Standard Scores <70 considered impaired, scores >70 and <85 considered below average

[b] Scaled Scores <4 considered impaired, scores >4 and <7 considered below average

[c] T-Scores>60 considered clinically significant

References

1. Gurney JG, Severson RK, Davis S, et al. Incidence of cancer in children in the United States: sex, race, and 1 year age-specific rates. Cancer. 1995; 75:2186–2195. [PubMed: 7697611]
2. Ganyon PS, Angiolillo AL, Carroll WL, et al. Long term results of the Children's Cancer Group studies for childhood acute lymphoblastic leukemia 1983-2002: A Children's Oncology Group Report. Leukemia. 2010; 24:285–297. [PubMed: 20016531]
3. Salzer WL, Devidas M, Carroll WL, et al. Long-term results of the Pediatric Oncology Group studies for childhood acute lymphoblastic leukemia 1984-2001: a report from the Children's Oncology Group. Leukemia. 2010; 24:355–370. [PubMed: 20016527]
4. Peterson CC, Johnson CE, Ramirez LY, et al. A meta-analysis of the neuropsychological sequelae of chemotherapy-only treatment for pediatric acute lymphoblastic leukemia. Pediatr Blood Cancer. 2008; 51:99–104. [PubMed: 18322925]
5. 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. 2011; 4:1–12. [PubMed: 21211043]
6. Shuper A, Stark B, Kornreich L, et al. Methotrexate- related neurotoxicity in the treatment of childhood acute lymphoblastic leukemia. IMAJ. 2002; 4:1050–1053. [PubMed: 12489505]
7. Peylan-Ramu N, Poplack DG, Pizzo PA, et al. Abnormal CT scans of the brain in asymptomatic children with acute lymphocytic leukemia after prophylactic treatment of the central nervous system with radiation and intrathecal chemotherapy. N Engl J Med. 1978; 298:815–819. [PubMed: 273143]
8. Ochs JJ, Berger P, Brecher ML, et al. Computed tomography brain scans in children with acute lymphocytic leukemia receiving methotrexate alone as central nervous system prophylaxis. Cancer. 1980; 45:2274–2278. [PubMed: 6929724]

9. Brown RT, Madan-Swain A, Walco GA, et al. Cognitive and academic late effects among children previously treated for acute lymphocytic leukemia receiving chemotherapy as CNS prophylaxis. J Pediatr Psychol. 1998; 23:333–340. [PubMed: 9782681]
10. Mahoney DH, Shuster JJ, Nitschke R, et al. Acute neurotoxicity in children with B-precursor acute lymphoid leukemia: an association with intermediate-dose intravenous methotrexate and intrathecal triple therapy-a Pediatric Oncology Group study. J Clin Oncol. 1998; 16:1712–1722. [PubMed: 9586883]
11. Chauvenet AR, Martin PL, Devidas M, et al. Anti-metabolite therapy for lesser risk B-lineage acute lymphoblastic leukemia of childhood: a report from Children's Oncology Group Study P9201. Blood. 2007; 110:1105–1111. [PubMed: 17442849]
12. Reddick WE, Glass JO, Helton KJ, et al. Prevalence of leukoencephalopathy in children treated for acute lymphoblastic leukemia with high-dose methotrexate. AJNR. 2005; 26:1263–1269. [PubMed: 15891195]
13. Chu WCW, Chik K, Chan Y, et al. White matter and cerebral metabolite changes in children undergoing treatment for acute lymphoblastic leukemia: longitudinal study with MR imaging and 1 H MR spectroscopy. Radiology. 2003; 229:650–669.
14. Wilson DA, Nitschke R, Bowman ME, et al. Transient white matter changes on MR images in children undergoing chemotherapy for acute lymphocytic leukemia: Correlation with neuropsychologic deficiencies. Radiology. 1991; 180:205–209. [PubMed: 2052695]
15. Montour-Proulx I, Kuehn SM, Keene DL, et al. Cognitive changes in children treated for acute lymphoblastic leukemia with chemotherapy only according to the Pediatric Oncology Group 9605 protocol. J Child Neurol. 2005; 20:129–133. [PubMed: 15794179]
16. Anderson FS, Kunin-Batson AS. Neurocognitive late effects of chemotherapy in children: The past 10 years of research on brain structure and function. Pediatr Blood Cancer. 2009; 52:159–164. [PubMed: 18680151]
17. Buizer AI, de Sonneville LMJ, Veerman AJP. Effects of chemotherapy on neurocognitive function in children with acute lymphoblastic leukemia: A critical review of the literature. Pediatr Blood Cancer. 2009; 52:447–454. [PubMed: 19061221]
18. Campbell LK, Scaduto M, Sharp W, et al. A meta-analysis of the neurocognitive sequelae of treatment for childhood acute lymphocytic leukemia. Pediatr Blood Cancer. 2007; 49:65–73. [PubMed: 16628558]
19. Kadan-Lottick NS, Brouwers P, Breiger D, et al. A comparison of neurocognitive functioning in children previously randomized to dexamethasone or prednisone in the treatment of childhood acute lymphoblastic leukemia. Blood. 2009; 114:1746–1752. [PubMed: 19546477]
20. Lofstad GE, Reinfjell T, Hestad K, et al. Cognitive outcome in children and adolescents treated for acute lymphoblastic leukemia with chemotherapy only. Acta Paediatr. 2009; 98:180–186. [PubMed: 18826490]
21. Mennes M, Stiers P, Vandenbussche E, et al. Attention and information processing in survivors of childhood acute lymphoblastic leukemia treated with chemotherapy only. Pediatr Blood Cancer. 2005; 44:479–486.
22. Buizer AI, de Sonneville LMJ, 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. 2005; 45:281–290. [PubMed: 15806539]
23. Kadan-Lottick NS, Brouwers P, Breiger D, et al. Comparison of neuro-cognitive functioning in children previously randomly assigned to intrathecal methotrexate compared with triple intrathecal therapy for the treatment of acute lymphoblastic leukemia. J Clin Oncol. 2009; 27:5986–5992. [PubMed: 19884541]
24. Cole PD, Kamen BA. Delayed neurotoxicity associated with therapy for children with acute lymphoblastic leukemia. MRDD Research Reviews. 2006; 12:174–183.
25. Ashley DM, Merchant TE, Strother D, et al. Induction chemotherapy and conformal radiation for very young children with non-metastatic medulloblastoma: P9934 A Children's Oncology Group study. J Clin Oncol. 2012; 30:3181–3186. [PubMed: 22851568]
26. Waber DP. More good news about neuropsychological late effects in long-term survivors of acute lymphoblastic leukemia. J Pediatr Hematol Oncol. 2002; 24:86–87. [PubMed: 11990710]

Table 1

Intensive CNS directed therapy during consolidation (P9605/P9201)

WEEK	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
P9201/P9605			IV			IV			IV			IV			IV			IV			
P9201/P9605			LV			LV			LV			LV			LV			LV			
P9201	IT*			IT		IT			IT			IT			IT			IT			IT
P9605	IT**	IT			IT			IT			IT			IT							IT

IV=intravenous methotrexate

LV= leucovorin rescue

* P9201 IT=intrathecal methotrexate (24/28 patients, 4/28 received TIT)

** P9605 IT=17/38 patients received TIT, 21/38 received varying combinations of TIT and ITM (see text)

J Pediatr Hematol Oncol. Author manuscript; available in PMC 2015 June 14.

Table 2

Demographics

	P9201 (n=24)	P9605 (n=35)	P value
Gender (n/%)			0.43
male	15 (62.5%)	17 (48.6%)	
female	9 (37.5%)	28 (51.4%)	
Race (n/%)			0.16
White, non-hispanic	16 (66.7%)	31 (88.6%)	
Black, non-hispanic	1 (4.2%)	2 (5.7%)	
Filipino	1 (4.2%)	0 (0%)	
Native American	1 (4.2%)	0 (0%)	
Hispanic	5 (20.8%)	2 (5.7%)	
Age at diagnosis*	4.1 (1.1-7.5)	4.9 (2.3-9.8)	0.09
Diagnosis to NP test*^	8.0 (6.1-9.6)	7.7 (5.2-10.9)	0.48
Diagnosis to MRI*^	7.7 (5.1-9.6)	7.8 (5.2-10.2)	0.74
End of treatment to NP test*^	5.6 (3.5-7.1)	5.2 (2.6-8.3)	0.25
End of treatment to MRI*^	5.3 (2.6-7.1)	5.3 (2.7-7.7)	0.94

*mean/range in years

^Only include patients with the particular test, as not all patients had both.

Table 3

Mean scores on P9605 and P9201 for each neurocognitive measure and P value for comparing the mean score between the two studies.

Variable	P9201		P9605		P value
	N	Mean	N	Mean	
WISC-3 Verbal IQ (VIQ) [a]	16	101.6	31	93.5	0.13
WISC-3 Performance IQ (PIQ) [a]	16	98.3	31	90.3	0.15
WISC-3 Full Scale IQ (FSIQ) [a]	17	100.4	32	90.1	0.06
WISC-3 Verbal Comprehension Index (VCI) [a]	17	102.9	31	93.9	0.07
WISC-3 Perceptual Organization Index (POI) [a]	16	102.1	27	90.7	**0.03**
WISC-3 Freedom from Distractibility Index (FDI) [a]	16	95.9	28	93.6	0.66
WISC-3 Processing Speed Index (PSI) [a]	15	98.1	26	89.8	0.14
WRAML Picture Memory Subtest [b]	17	10.4	33	9.5	0.33
WRAML Design Memory Subtest [b]	18	9.3	32	9.2	0.91
WRAML Verbal Learning Subtest [b]	18	10.7	33	9.9	0.51
WRAML Story Memory Subtest [b]	18	10.7	32	9.7	0.24
WRAML Overall Memory Index [a]	17	101.3	32	97.6	0.49
WJ-R Visual-Matching Subtest [a]	14	97.9	29	87.7	0.06
WJ-R Cross Out Subtest [a]	14	102.8	29	91.5	0.06
WJ-R Processing Speed Cluster Index [a]	14	99.9	28	88.1	**0.05**
NEPSY Tower Subtest [b]	12	16.4	17	10.2	0.25
Beery VMI [a]	17	104.6	32	92.6	**0.03**
CPT-II Hits [c]	7	85.3	8	77.4	0.59
CPT-II Omissions [c]	17	65.5	32	58.1	0.43
CPT-II Commissions [c]	17	52.9	33	46.5	**0.05**
CPT-II Hit Rt [c]	16	45.0	33	49.2	0.32
CPT-II Hit RT SE [c]	17	55.7	33	54.7	0.82
CPT-II Variability [c]	17	52.9	33	52.7	0.96
CPT-II Attentiveness [c]	10	56.4	23	52.3	0.46
CPT-II Risk [c]	10	61.9	23	58.0	0.65
CPT-II Hit Rt Block Change [c]	17	53.7	33	53.0	0.87
CPT-II Hit SE Block Change [c]	17	50.3	33	52.7	0.51

Variable	P9201		P9605		P value
	N	Mean	N	Mean	
CPT-II Hit Rt ISI Change [c]	17	53.9	33	54.5	0.91
CPT-II Hit SE ISI Change[c]	17	50.1	33	51.6	0.71

[a] Standard Score (M=100, SD=15)

[b] Scaled Score (M=10, SD=3)

[c] T-Score (M=50, SD=10)

Table 4

Proportion of patients with neurocognitive subtest scores 1 or more standard deviation worse than the normative mean, overall and stratified by study, and P value for comparing the observed % to 16%.

	P9201		P9605		Overall	
	%	P value	%	P value	%	P value
WISC-3 Verbal IQ (VIQ) [a]	12.5%	1.00	29.0%	0.10	23.4%	0.24
WISC-3 Performance IQ (PIQ) [a]	18.8%	0.97	38.7%	**0.004**	31.9%	**0.01**
WISC-3 Full Scale IQ (FSIQ) [a]	5.9%	0.44	31.3%	**0.05**	22.4%	0.30
WISC-3 Verbal Comprehension Index (VCI) [a]	11.8%	0.95	29.0%	0.10	22.9%	0.27
WISC-3 Perceptual Organization Index (POI) [a]	6.3%	0.50	25.9%	0.26	18.6%	0.76
WISC-3 Freedom from Distractibility Index (FDI) [a]	31.3%	0.20	25.0%	0.30	27.3%	0.08
WISC-3 Processing Speed Index (PSI) [a]	13.3%	1.00	38.5%	**0.01**	29.3%	**0.05**
WRAML Picture Memory Subtest [b]	11.8%	0.95	15.2%	1.00	14.0%	0.88
WRAML Design Memory Subtest [b]	11.1%	0.87	25.0%	0.26	20.0%	0.54
WRAML Verbal Learning Subtest [b]	11.1%	0.87	18.2%	0.87	15.7%	1.00
WRAML Story Memory Subtest [b]	11.1%	0.87	12.5%	0.81	12.0%	0.58
WRAML Overall Memory Index [a]	11.8%	0.95	21.9%	0.49	18.4%	0.77
WJ-R Visual-Matching Subtest [a]	21.4%	0.79	31.0%	0.07	27.9%	0.07
WJ-R Cross Out Subtest [a]	0%	**<0.001**	24.1%	0.34	16.3%	1.00
WJ-R Processing Speed Cluster Index [a]	21.4%	0.79	32.1%	**0.05**	28.6%	0.06
NEPSY Tower Subtest [b]	8.3%	0.81	5.9%	0.44	6.9%	0.27
Beery VMI [a]	5.9%	0.44	28.1%	0.12	20.4%	0.50
CPT-II Hits [c]	85.7%	**<0.001**	75.0%	**0.001**	80.0%	**<0.001**
CPT-II Omissions [c]	52.9%	**0.001**	40.6%	**0.002**	44.9%	**<0.001**
CPT-II Commissions [c]	35.3%	0.08	12.1%	0.75	20.0%	0.54
CPT-II Hit Rt [c]	12.5%	1.00	15.2%	1.00	14.3%	0.93
CPT-II Hit RT SE [c]	29.4%	0.24	30.3%	0.06	30.0%	**0.02**
CPT-II Variability [c]	41.2%	**0.02**	30.3%	0.06	34.0%	**0.003**
CPT-II Attentiveness [c]	60.0%	**0.004**	17.4%	1.00	30.3%	0.06
CPT-II Risk [c]	50.0%	**0.03**	30.4%	0.12	36.4%	**0.01**
CPT-II Hit Rt Block Change [c]	29.4%	0.24	21.2%	0.54	24.0%	0.19
CPT-II Hit SE Block Change [c]	23.5%	0.57	21.2%	0.54	22.0%	0.33

	P9201		P9605		Overall	
	%	P value	%	P value	%	P value
CPT-II Hit Rt ISI Change [c]	17.6%	1.00	27.3%	0.14	24.0%	0.19
CPT-II Hit SE ISI Change[c]	29.4%	0.24	24.2%	0.29	26.0%	0.10

[a] Standard Score (M=100, SD=15)

[b] Scaled Score (M=10, SD=3)

[c] T-Score (M=50, SD=10)