Reduced Frontal White Matter Volume in Long-Term Childhood Leukemia Survivors: A Voxel-Based Morphometry Study

M.E. Carey, M.W. Haut, S.L. Reminger, J.J. Hutter, R. Theilmann and K.L. Kaemingk

AJNR Am J Neuroradiol 2008, 29 (4) 792-797
doi: https://doi.org/10.3174/ajnr.A0904
http://www.ajnr.org/content/29/4/792
Reduced Frontal White Matter Volume in Long-Term Childhood Leukemia Survivors: A Voxel-Based Morphometry Study

BACKGROUND AND PURPOSE: To our knowledge, no published studies have examined whole-brain regional differences to identify more discrete volumetric changes in the brains of childhood leukemia survivors. We used voxel-based morphometry (VBM) to examine regional gray and white matter differences in a group of long-term survivors of acute lymphoblastic leukemia (ALL) compared with a group of healthy controls. Differences in regional white matter volume were expected, given previous reports of white matter changes during treatment for ALL and reduced brain white matter volumes in long-term survivors. Follow-up analyses examined the relationship of regional brain volumes to cognitive function.

MATERIALS AND METHODS: We compared 9 long-term survivors of ALL with 14 healthy controls. Survivors of ALL were treated with systemic and intrathecal chemotherapy only. T1-weighted axial 3D spoiled gradient high-resolution images collected on a 1.5T MR imaging scanner were used for the VBM analysis. Neuropsychological evaluations were conducted within 2 months of the MR imaging to assess cognitive function.

RESULTS: VBM analysis revealed 2 specific regions of reduced white matter in the right frontal lobes of survivors of ALL compared with healthy controls. Survivors of ALL had lower performances on tests of attention, visual-constructional skills, mental flexibility, and math achievement compared with healthy individuals. Decreased performance on neuropsychological measures was associated with decreased regional white matter volumes. No differences were found between the groups with respect to gray matter regions.

CONCLUSION: These findings are consistent with previous literature describing the long-term cognitive, academic, and imaging findings of survivors of ALL and suggest that right frontal white matter is particularly vulnerable to disruption following intensive chemotherapy for ALL. Future studies should focus on further clarifying the white matter changes observed.

Acute lymphoblastic leukemia (ALL) is the most common cause of childhood cancer in the United States. Currently, most children treated for ALL receive central nervous system (CNS) prophylaxis with a combination of intrathecal and systemic chemotherapy. Intrathecal chemotherapy protocols were developed, in part, because of the known toxic effect of cranial radiation therapy and its adverse impact on cognitive function. CNS chemotherapy has been very effective, with the incidence of CNS relapse reduced from 75% to 80% to less than 10%. Unfortunately, there is accumulating evidence that treatment protocols for ALL have acute and long-term effects on the brain, despite the elimination of cranial radiation therapy from standard treatment.

CNS structural abnormalities have consistently been reported in neuroimaging studies of children with ALL treated with prophylactic intrathecal chemotherapy without radiation therapy. The most frequent findings are white matter changes, widening of the ventricles and/or sulci, and cerebral calcifications. Reduced prefrontal cortex, cerebellum, and cortical white matter volumes have also been reported. In addition, long-term difficulties in cognitive and academic function are often described, most frequently visual perceptual, attention, and math skills.

Voxel-based morphometry (VBM) is an automated method for analyzing whole brain regional differences on a voxel-by-voxel basis. The purpose of this study was to use VBM to examine regional white and gray matter differences in a group of long-term survivors of ALL treated with systemic and intrathecal chemotherapy, compared with a group of healthy controls, and to examine the relationship of regional brain volumes to cognitive function. To our knowledge, no published studies have examined whole-brain regional differences to identify more discrete volumetric changes in the gray and white matter of survivors of childhood leukemia. Given previous reports of white matter changes during treatment for ALL and reduced brain white matter volumes in long-term survivors, we hypothesized that survivors of ALL would have reduced regional white matter volumes and impaired cognitive function compared with healthy controls.

Methods

Participants

All survivors of childhood ALL, who were between 6 and 30 years of age and were diagnosed before age 18, were recruited to participate in...
this study through the University of Arizona (UA) Pediatric Hematology/Oncology Late Effects Clinic between June 2003 and June 2005. All survivors of childhood ALL followed at the UA Late Effects Clinic, or their parents/legal guardians, were informed of the study in a letter. Additional participants were recruited through study brochures available in the UA Pediatric Outpatient Clinic and through a local press release about the study, both of which were institutional review board–approved. Potential healthy control participants were primarily recruited from the siblings of participating survivors of ALL.

Leukemia survivors in this study were diagnosed between 1987 and 2000. Therefore, specific treatment protocols varied across participants. However, all survivors received CNS prophylaxis with intrathecal chemotherapy during treatment. None of the survivors received cranial radiation therapy or bone marrow transplant, and none had a history of recurrence. Exclusionary criteria were non-English-speaking individuals, mental retardation, psychiatric disorder, neurologic disorder, traumatic brain injury associated with an alteration of consciousness, and premorbid diagnoses of learning disability or attention deficit–hyperactivity disorder. Only those participants who could complete the MR imaging without sedation and who did not have metallic objects present in their bodies were included in the MR imaging examination. This study was approved by the UA institutional review board committee. Informed consent was obtained from all parents or participants of legal consenting age, and assent for participation was obtained for all minor-aged children. Participants received $25 for completion of the cognitive assessment and $25 for completion of the MR imaging.

A total of 14 survivors of ALL and 18 healthy controls enrolled in the study. Thirteen survivors of ALL and 17 healthy controls completed both the MR imaging and the neuropsychological evaluation. The MR images of 4 survivors of ALL and 3 healthy controls were not usable due to movement artifact or distortion. Therefore, we conducted analyses with 9 survivors of ALL and 14 healthy participants. There was no difference in mean age between the ALL group (mean = 15.17 years, SD = 5.48, range = 7.75–25.76) and the healthy control group (mean = 15.22, SD = 3.78, range = 8.43–21.51). The ALL group consisted of 3 males and 6 females, all of whom were right-handed. Mean age at diagnosis was 5.17 years (SD = 2.96, range = 1.43–9.36). Mean time since diagnosis was 9.95 years (SD = 5.13, range = 3.48–16.96). The healthy control group consisted of 7 males and 7 females, 3 of whom were left-handed. The average number of days between the MR imaging and the neuropsychological evaluation was 11.30 (range = 0–60 days). Nonparametric analyses did not reveal any significant differences between the groups with respect to race/ethnicity, parent education, or household income.

**MR Imaging Acquisition**

Structural MR imaging data were collected on a 1.5T scanner (GE Healthcare, Milwaukee, Wis) at the University Medical Center. The scans included a 3D T1-weighted localizer scan. This scan consisted of 93 (31 in each plane) 4-mm images by using a fast spin-echo pulse sequence with the following parameters: 256 × 128 matrix, 1 signal-intensity average, 90° flip angle, minimal TE and TR, FOV = 24, scanning time = 2 minutes 4 seconds. There were also 2 simultaneous series (depending on head size) of 58 to 68 three-millimeter images collected using an oblique dual fast spin-echo pulse sequence (T2- and proton density–weighted scans) in a plane perpendicular to the long axis of the hippocampus with the following parameters: 256 × 256 matrix; 90° flip angle; TE = 17,102 ms and TR = 3800 ms; FOV = 24; scanning time = 12 minutes 33 seconds. The last was a series of 124 T1-weighted axial 3D spoiled gradient-recalled 1.5-mm high-resolution images with the following parameters: 256 × 256 matrix, 30° flip angle, TE = 5 ms, TR = 22 ms, FOV = 25, scanning time = 12 minutes 4 seconds. The spoiled gradient-recalled (SPGR) images were used in the current analyses.

**VBM and Regional Analysis**

VBM was performed by using the optimized method implemented with Statistical Parametrical Mapping-2 (SPM2; Wellcome Department of Imaging Neuroscience, London, UK; http://www.fil.ion.ucl.ac.uk/spm/). First, a local sample-specific template was created due to developmental differences in pediatric compared with adult populations and potential problems with adult a priori information during spatial normalization and segmentation. The local, sample-specific template was created by first normalizing each individual SPGR image from our entire sample to the standard SPM2 T1 template. The standard automated normalization routine in SPM2 uses a combination of linear and nonlinear methods. The first step in normalization involves a 12-parameter affine transformation by using a Bayesian method to search for the best match. Then nonlinear deformations are performed by using a linear combination of discrete cosine transform basis functions in 3 orthogonal directions. Each volume was resampled into 2-mm³ isotropic voxels. These normalized volumes were averaged and smoothed with a 12-mm full width half maximum filter to produce the local sample-specific template. Then each individual SPGR image was segmented into gray, white, and CSF volumes by using the local sample-specific template for a priori information. The segmentation routine is an automated procedure that uses a priori information about the likelihood that voxels are a specific tissue type (eg, gray matter) within a mixture model cluster analysis.

The segmented gray and white matter volumes were then placed into standard space using the local sample-specific gray and white matter templates, created by segmenting the local sample-specific template, with the combination of linear and nonlinear methods detailed previously. Each volume was corrected for the effects of non-linear normalization by multiplying each voxel by the Jacobian determinants derived during the spatial normalization step. The normalized volumes were resampled into 2-mm³ isotropic voxels and smoothed with a 12-mm full width half maximum filter to account for individual differences in anatomy. Using the general linear model, a voxel-by-voxel search covarying for full-scale intelligence quotient (FSIQ) was then conducted to identify regions that were different between groups, with P set at <.001, uncorrected. The minimal number of contiguous voxels was set at 30.

**Neuropsychologic Measures**

Neuropsychologic variables used are listed in Table 1 and include tests of intelligence, visual-constructional skills, language, academic achievement, memory, attention, processing speed, and executive function. All statistical analyses were conducted by using the Statistical Package for the Social Sciences 15.0 for Windows (1989–2006; SPSS, Chicago, Ill). Multivariate analysis of covariance (MANCOVA) was used for between-group comparisons on cognitive variables. Although the differences between the groups with respect to mean FSIQ did not reach statistical significance (P = .059), the >10-point mean FSIQ difference between the groups does have a clinical significance (Table 2). Therefore, FSIQ was entered as a covariate for the MANCOVA. Follow-up 1-tailed Pearson correlational analyses were conducted by using the whole sample (survivors of ALL and healthy con-
analyses. Due to the small sample size, a significance level of $P < .01$ was used for all analyses of neuropsychological data to reduce the chance of a type I error.

Results

VBM and Regional Analysis

Regional analysis revealed decreased white matter volume in the regions of the right middle frontal gyrus ($x, y, z = 32, 38, 36$; $t = 4.94$; voxels $= 42$; Fig 1A, -B) and the right superior frontal gyrus ($x, y, z = 16, -4, 64$; $t = 4.52$; voxels $= 62$; Fig 2A, -B). In both cases, white matter volumes were reduced in the survivors of ALL compared with healthy controls. No regions were identified in which healthy controls had decreased volumes compared with survivors of ALL. There were no differences between the groups with respect to regional gray matter volumes in either direction.

Cognitive Variables Analyses

The ALL group had significantly lower mean scores on the Wechsler Block Design ($P = .001$), Vocabulary ($P < .001$), and Digit Span ($P < .001$), as well as the Delis–Kaplan Executive Function System Trail Making 4 ($P = .002$), and Verbal Fluency Category Switching ($P = .007$) even after covarying for FSIQ. A nonsignificant trend for a lower mean score on Woodcock-Johnson III Tests of Achievement Calculation ($P = .014$) in the ALL group was found as well. See Table 2 for group score comparisons.

One-tailed Pearson correlations revealed a significant relationship between white matter volume in the region of the right middle frontal gyrus and Woodcock-Johnson III Tests of Achievement Calculation ($r = 0.651, P < .001$; Fig 3), with decreased volume associated with decreased math performance. Nonsignificant but strong correlations between this area and Trail Making 4 ($r = 0.486, P = .011$) and Digit Span ($r = 0.420, P = .023$) were also found, with decreased volume associated with decreased performance. One-tailed Pearson correlations also revealed a significant relationship between white matter volume in the region of the right superior frontal gyrus and Vocabulary ($r = 0.505, P = .007$; Fig 4), with decreased volume associated with decreased performance. Nonsignificant but strong correlations between this area and Digit Span ($r = 0.436, P = .019$) as well as Block Design ($r = 0.427, P = .021$) were also found, with decreased volume associated with decreased performance.

Discussion

Consistent with previous research,6-9,11,12,16,18,20 we found evidence of white matter brain abnormalities in survivors of ALL. To our knowledge, this is the first study that used VBM techniques to identify specific regions of decreased brain volume in childhood leukemia survivors. In particular, we found 2 areas of decreased regional white matter volume in the areas of the right middle and right superior frontal gyri of long-term survivors of ALL compared with healthy controls.

Reduced white matter volume in survivors of ALL has previously been reported.7 Focusing on a transverse region of interest covering most of the corpus callosum, Reddick et al reported that 84 survivors of ALL (ages 6–18 years) treated with chemotherapy alone had significantly smaller white mat-

### Table 1: Neuropsychological measures used

<table>
<thead>
<tr>
<th>Test</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wechsler Full Scale IQ</td>
<td>The WISC-III was used for ages 6–18; the WAIS-III was used for ages 17 and old</td>
</tr>
<tr>
<td>Wechsler Block Design</td>
<td>A visual-constructural task using colored blocks</td>
</tr>
<tr>
<td>Wechsler Vocabulary</td>
<td>A language task that requires the subject to provide word definitions</td>
</tr>
<tr>
<td>Wechsler Digit Span</td>
<td>A test of attention/working memory requiring repetition of number sequences in forward and reverse order</td>
</tr>
<tr>
<td>Wechsler Coding</td>
<td>This is a 2-minute timed processing-speed task that requires the individual to copy symbols quickly</td>
</tr>
<tr>
<td>WJ-III Letter-Word ID</td>
<td>A measure of letter identification and single-word reading</td>
</tr>
<tr>
<td>WJ-III Calculation</td>
<td>A measure of written mathematical computation skills</td>
</tr>
<tr>
<td>WJ-III Story Recall</td>
<td>A measure of immediate memory for increasingly complex stories that are presented using an audio recording</td>
</tr>
</tbody>
</table>

### Table 2: Group comparisons on neuropsychological measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>ALL Mean (SD)</th>
<th>Control Mean (SD)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wechsler FSIQ (mean, 100, SD, 15)</td>
<td>101.67 (8.28)</td>
<td>113.00 (15.62)</td>
<td>.059</td>
</tr>
<tr>
<td>Block Design (mean, 10, SD, 3)</td>
<td>11.22 (2.17)</td>
<td>12.71 (3.69)</td>
<td>.001</td>
</tr>
<tr>
<td>Vocabulary (mean, 10, SD, 3)</td>
<td>10.44 (2.70)</td>
<td>12.00 (3.19)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Coding (mean, 10, SD, 3)</td>
<td>11.56 (3.32)</td>
<td>10.79 (3.07)</td>
<td>NS</td>
</tr>
<tr>
<td>Digit Span (mean, 10, SD, 3)</td>
<td>8.67 (3.08)</td>
<td>11.71 (1.12)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>WJ-III Letter-Word ID</td>
<td>100.11 (11.25)</td>
<td>105.07 (11.09)</td>
<td>NS</td>
</tr>
<tr>
<td>Calculation</td>
<td>104.22 (13.81)</td>
<td>115.07 (14.43)</td>
<td>.014</td>
</tr>
<tr>
<td>Story Recall</td>
<td>100.33 (13.76)</td>
<td>108.36 (11.70)</td>
<td>NS</td>
</tr>
<tr>
<td>D-KEFS (mean, 10, SD, 3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trail Making Test 4</td>
<td>9.25 (2.55)</td>
<td>10.57 (3.28)</td>
<td>.002</td>
</tr>
<tr>
<td>Letter Fluency</td>
<td>10.88 (4.22)</td>
<td>10.86 (3.84)</td>
<td>NS</td>
</tr>
<tr>
<td>Category Fluency</td>
<td>11.88 (4.22)</td>
<td>11.79 (3.66)</td>
<td>NS</td>
</tr>
<tr>
<td>Category Switching-correct responses</td>
<td>10.88 (2.70)</td>
<td>11.14 (3.59)</td>
<td>.007</td>
</tr>
<tr>
<td>Tower Test-total achievement</td>
<td>10.88 (1.55)</td>
<td>10.93 (2.13)</td>
<td>NS</td>
</tr>
</tbody>
</table>

ter volumes compared with 33 healthy siblings (ages, 6–16 years). Reduced white matter volume was related to greater deficits on measures of attention, intelligence, and academic achievement. Wilson et al. investigated white matter changes on MR imaging in 25 children with ALL during treatment with systemic (intravenous) and intrathecal chemotherapy. They reported that 17 of the 25 (68%) children showed transient white matter hyperintensities during the first year of treatment. Furthermore, these white matter changes closely paralleled the administration of intravenous methotrexate (MTX), possibly in conjunction with intrathecal chemotherapy. In a longitudinal study of survivors of ALL, Harila-Saari et al. reported that 3 of 30 patients (10%) had MR imaging white matter hyperintensities at the completion of treatment and 4

![Image](https://example.com/image1)

**Fig 1.** T1-weighted axial (A) and sagittal spoiled gradient-recalled (B) images show the right middle frontal gyrus with healthy controls versus survivors of ALL with respect to regional white matter volume. Note that images have been left-right reversed for presentation.

![Image](https://example.com/image2)

**Fig 2.** T1-weighted axial (A) and sagittal spoiled gradient-recalled (B) images show the right superior frontal gyrus with healthy controls versus survivors of ALL with respect to regional white matter volume. Note that images have been left-right reversed for presentation.

![Image](https://example.com/image3)

**Fig 3.** Correlation between the right middle frontal gyrus region of interest (ROI) (cubic millimeters) and Calculation.

![Image](https://example.com/image4)

**Fig 4.** Correlation between the right superior frontal gyrus region of interest (ROI) (cubic millimeters) and Vocabulary.
of 32 patients (12.5%) had white matter hyperintensities 5 years posttreatment. Asato et al6 followed 16 children diagnosed with ALL and 4 diagnosed with malignant lymphoma in a prospective study by using serial MR images before and during the induction-consolidation phases of treatment. Areas of high signal intensity were found in the cerebral deep white matter of 8 children. Seven of the 8 children showed increased signal intensity areas over the serial MR images. These changes were transient in 6 of the 8 children. The investigators further reported that these areas of high signal intensity were first observed following the introduction of MTX between weeks 5 and 29 and that significant improvements were only noted on MR imaging after the interruption or discontinuation of intrathecal MTX.

The mean age of diagnosis for ALL is 5 years. Brain myelination continues through adolescence and occurs in an inferior-to-superior and posterior-to-anterior fashion.32 For these reasons, it is reasonable to suspect that early exposure to toxic chemotherapy agents in the CNS may disrupt development of frontal lobe white matter. Our finding of reduced regional white matter volumes in the frontal lobes of survivors of childhood ALL supports this idea and is consistent with a previous report of reduced prefrontal cortex volumes among long-term survivors of ALL.13 Lesnik et al13 found reduced prefrontal cortex and cerebellar volumes in 10 long-term survivors of ALL treated only with chemotherapy when compared with 10 healthy controls matched for age and sex. The volume differences were related to performances on measures of attention, memory, and executive function.

Frontal lobe abnormalities are associated with impairments in attention and executive function skills,43 and attention difficulties are frequently reported in long-term survivors of ALL.7,44 In the current study, verbal (Digit Span) and visual (Trail Making) attention measures were significantly weaker for survivors of ALL compared with the healthy controls, and these measures correlated \((r > 0.420)\) with regional frontal white matter volumes, though the correlations did not reach statistical significance by using a \(P\) value of <.01 in this small sample.

Math achievement is the most frequently reported area of academic difficulty among survivors of ALL.26,33-35 Consistent with this report, written calculation skills were significantly lower in our group of survivors of ALL compared with healthy controls, and performance was significantly related \((r = 0.651)\) to white matter volume in the area of the right middle frontal gyrus. Although math skills, particularly mental arithmetic skills, are most often associated with intraparietal sulcus activation during imaging studies,45-47 functional imaging studies have also found bilateral prefrontal cortex activation during numeric processing and mental arithmetic tasks in children and adults.45,47 Furthermore, a recent study of healthy children diagnosed with developmental dyscalculia reported reduced brain activation in the right middle frontal gyrus as well as in the left intraparietal sulcus and left inferior frontal gyrus of these children compared with healthy controls.46 Among childhood brain tumor survivors, the volume of normal-appearing white matter, in combination with attentional abilities and IQ, has been found to account for a significant proportion of the variance in math achievement.48 Therefore, reduced math achievement among survivors of ALL in this study might also be explained by reduced attention, IQ, and regional frontal lobe white matter volume.

The finding that vocabulary was significantly related to the volume of white matter in the area of the right superior frontal gyrus was not expected, given the well-established relationship of language function with the left hemisphere in most normally developing people.43 However, vocabulary is often used as a measure of general intelligence, because it is highly correlated with IQ and is not a cognitive skill typically associated with a focal brain region. This finding may, therefore, be more indicative of the known relationship of global intellectual function with overall white matter volume.49,50

Survivors of ALL in our study had significant difficulty on tests of attention, visual-constructional skills, mental flexibility, and math achievement compared with healthy individuals. This neuropsychological profile is commonly reported in long-term survivors of ALL and is sometimes described as a right hemisphere or nonverbal learning disability.51 Most interesting, performance on these neuropsychological measures was related to white matter volume of the identified regions of interest. These findings further support the notion that cognitive difficulties experienced by survivors of ALL is secondary to changes in white matter development,7 particularly of the right hemisphere,51 and validate the current MR imaging findings.

This pilot study was limited by a small sample size. Despite this limitation, the results were consistent with previous literature describing the long-term cognitive, academic, and imaging findings of survivors of ALL and suggest that the right frontal white matter is particularly vulnerable to disruption following intensive chemotherapy for childhood leukemia. Future studies should focus on further clarifying the white matter changes observed. In particular, diffusion tensor imaging may shed some light on the more subtle white matter changes occurring in normal-appearing white matter. Diffusion tensor imaging could also be useful in determining if the integrity of white matter is disrupted throughout the brain or if it is truly circumscribed to the frontal lobe areas of survivors of ALL. This would lend further support to the idea that chemotherapy at a young age is disruptive to myelination of the frontal lobes in survivors of ALL. In addition, functional neuroimaging studies may provide insight into how early disruption of white matter development impacts the functional brain organization of survivors of ALL.

**Conclusion**

We used VBM to examine regional gray and white matter differences in a group of long-term survivors of ALL compared with a group of healthy controls. We identified 2 areas of decreased regional white matter volume in the areas of the right middle and right superior frontal gyri of survivors of ALL compared with healthy controls. Findings were consistent with previous evidence of white matter brain abnormalities and right hemisphere cognitive dysfunction following treatment for ALL. Future studies should focus on further clarifying the white matter changes observed.
Acknowledgment
We thank Joachim Seeger, MD, for his time and effort in reviewing the radiologic scans.

References