Voxel-Based Analysis of T2 Hyperintensities in White Matter during Treatment of Childhood Leukemia

BACKGROUND AND PURPOSE: White matter (WM) hyperintensities on T2-weighted MR imaging are the most common imaging manifestation of neurotoxic effects of therapy for central nervous system (CNS) prophylaxis in childhood acute lymphoblastic leukemia (ALL). This study uses voxel-based analyses (VBA) of T2-weighted imaging of patients during treatment to identify which WM regions are preferentially damaged.

MATERIALS AND METHODS: Two sets of conventional T2-weighted axial images were acquired on a 1.5T MR imaging scanner from 197 consecutive patients (85 female, 112 male; aged 1.0–18.9 years) enrolled on an institutional ALL treatment protocol. Images were acquired after completion of induction therapy and after the final of the 4 courses of intravenous high-dose methotrexate in consolidation therapy (3.9 ± 0.8 months apart). Voxel-wise statistical testing of the incremental change between normalized longitudinal T2 images was performed with radiologist reading (normal or abnormal) and treatment risk-group as covariates.

RESULTS: Two highly significant bilateral clusters of T2 signal intensity change were identified in both 1-group and 2-group analyses. The regions were symmetric in size, shape, and average signal intensity. Increased T2-weighted signal intensity from these regions both within and between examinations were nonlinear functions of age at examination, and the difference between the examinations was greater for older subjects who received more intense therapy.

CONCLUSIONS: These analyses identified specific WM tracts involving predominantly the anterior, superior, and posterior corona radiata and superior longitudinal fasciculus, which were at increased risk for the development of T2-weighted hyperintensities during therapy for childhood ALL. These vulnerable regions may be the cause of subsequent cognitive difficulties consistently observed in survivors.
were eligible for the study. ALL treatment was based on a comprehensive risk classification, including blast cell immunophenotype and genotype, presenting clinical features, and early treatment response. Patients were assigned to 1 of 3 risk groups: low-risk (LR) and standard-risk or high-risk (SHR). Patients between 1 and 10 years old with B-cell precursor and presenting leukocyte count < 50 × 10^9/L, leukemic cell DNA index of 1.16 or higher, or TEL-AML1 fusion were provisionally classified to have low-risk ALL, provided that they did not have testicular or CNS leukemia (ie, CNS3 status), hypodiploidy (< 45 chromosomes), E2A-PBX1 fusion, or MLL rearrangement. Patients with BCR-ABL fusion (Philadelphia chromosome) were designated to have high-risk disease, and all others including patients with T-cell ALL were provisionally classified to have standard-risk ALL. Final risk status depended on the response to remission-induction therapy. Any patients with 0.01% to 0.99% residual leukemia after completion of 6-week induction therapy were considered to have standard-risk ALL and received intensive postremission therapy, whereas those with 1% or more residual disease were designated to have high-risk ALL and were candidates for allogeneic hematopoietic stem cell transplantation.

During consolidation therapy, LR patients received IV-MTX at 2.5 g/m² plus intrathecal therapy every other week for 4 doses, whereas SHR patients received 5.0 g/m² together with intrathecal therapy. No patients received craniospinal irradiation. Written informed consent was obtained from the patient, parent, or guardian according to the Institutional Review Board, National Cancer Institute, and Office for Human Research Protections guidelines.

A single radiologist (F.H.L.) retrospectively reviewed imaging from the 287 subjects enrolled on the treatment protocol during the specified period who had completed therapy. The radiologist viewed all longitudinal imaging available from each patient throughout therapy and took into account the patients’ age at the time of examination. He used imaging from young healthy control subjects to assist in differentiating typical age-appropriate T2 hyperintensities in the terminal zones from those changes that were more likely therapy induced. Therapy-induced changes were assessed according to the Common Toxicity Criteria v2.0 for leukoencephalopathy, which requires focal or diffuse T2 hyperintensities involving the periventricular, centrum semiovale, or other susceptible WM areas of the cerebrum and an increase in subarachnoid space possibly with mild ventriculomegaly.

To be eligible for this study, patients had to complete MR imaging examinations at baseline after induction therapy and after the final course of IV-MTX in consolidation therapy. The differences in imaging between these 2 time points represent changes that occur during consolidation therapy. Eighty-four patients were missing at least 1 of the 2 examinations. Six additional patients with imaging findings of cerebral thrombosis, encephalomalacia, or large developmental abnormalities were also excluded from the study. These exclusion criteria resulted in a final cohort of 197 subjects (85 female; 112 male: 108 LR and 89 SHR aged 1.0–18.9 years [median, 5.3 years] at diagnosis. Time between the 2 examinations was 3.9 ± 0.8 months.

**MR Imaging**

Imaging was performed on 1 of 2 1.5T whole-body MR systems with use of the standard circular polarized volume head coil (Avanto; Siemens, Iselin, NJ). No hardware upgrades were performed during the conduct of this study. Software upgrades did occur during the study, but their impact on the routine T2-weighted imaging sequence used was negligible. Conventional imaging was acquired and included at least 19 4-mm-thick axial T1-weighted, T2-weighted, proton density (PD)-weighted, and fluid-attenuated inversion recovery (FLAIR)-weighted imaging sets, which covered most of the cerebrum starting at the apex of the brain but did not include the cerebellum because of limitations of coverage in the FLAIR sequence. The T2 imaging set has higher signal-to-noise intensity and less artifacts than the FLAIR imaging and was selected for this study of T2-hypertensities within the WM. Imaging was acquired as a T2-weighted, PD-weighted dual spin-echo image set (TR, 3500 ms; TE1, 17 ms; TE2, 102 ms; 7 echoes). Unfortunately, the imaging acquired in this treatment protocol, which began in 2000, did not include diffusion tensor imaging as a standard part of clinical imaging.

**Spatial Normalization**

VBA require that all examinations be registered into a common stereotactic space. For this study, each patient examination was registered to the ICBM T2 50 atlas (http://www.loni.ucla.edu/Atlases). This atlas is an average of the T2-weighted MR images of 50 healthy young adult brains in the deterministic atlas space of the ICBM 452 and not on any single subject. It represents the mean volume constructed from the average position, orientation, scale, and shear from all 50 individual subjects.

We achieved the spatial normalization of patient examinations to the T2 atlas using a free-form deformation (FFD) with B-spline interpolation, providing local support to model large focal warping included in the VTK CISG Registration Toolkit 2.0. To more closely match the T2 atlas, extrameningeal tissues were removed before registration with use of custom software on the basis of image intensity gradients. Default values for the FFD were used in all normalizations. The deformation field was determined by the displacements specified on a lattice of control points, and the displacements that were off the control points were interpolated on the basis of the neighborhood control points with use of uniform cubic B-splines. Normalized mutual information was used to measure image similarity during normalization with the FFD.

**VBA**

We performed voxel-wise statistical testing of the normalized longitudinal T2 images using the VBM Toolkit implemented within SPM2 (Wellcome Department of Imaging Neuroscience, London, UK; http://www.fil.ion.ucl.ac.uk/spm/). VBA is a widely used statistical image analysis approach for the detection of WM abnormalities that spatially normalizes (co-registers) brain images across subjects and performs statistical tests at each voxel. The advantages of VBA are that it is highly reproducible, user independent, and it can explore differences over the entire brain without anatomically specific hypotheses. In preparation for this analysis, a WM mask was created from the average of normalized T2-weighted images from the first examination with use of SPM2 to segment WM and then threshold the resulting image at a level of 0.7 probability to generate a binary WM mask. Second, the normalized T2 examinations were all smoothed with a full width half maximum 5-mm Gaussian kernel.

Longitudinal voxel-wise analyses were performed for 1 group with 2 conditions (examinations): after induction therapy and after the final course of IV-MTX. The radiologist reading of the examinations as normal (0) or abnormal (1) was included as a covariate. No nuisance variables were included, and a default 0.1 threshold was used in the analysis, along with the explicit mask of WM. Analyses were performed for all patients as a single group and then as a 2-group analysis stratified by treatment risk-group (LR and SHR). Age at the time of...
examination was not included in the analysis because it correlates highly with treatment risk-group.

**Statistical Analyses**

*F*-tests were performed to determine which voxels were significantly different between examinations over the subject populations. A *P* value and cluster threshold were specified to limit the analysis only to regions that have significant differences between examinations and have a sufficient number of continuous voxels for analysis. A false discovery rate approach, which controls the expected proportion of false-positives among suprathreshold voxels, was used to account for multiple comparisons. The results were overlaid onto the average of normalized T2-weighted images from the first examination for visualization.

**Results**

**One-Group Analysis**

The first analysis performed was for all 197 patients in the study regardless of risk stratification. Almost one third of the second examinations (57 [29%] of 197) were read by the radiologist as exhibiting T2 hyperintensities. It should be noted that 3 additional patients who tested normal on the early examinations went on to have T2 hyperintensities later in therapy. The large size of this sample enabled a quantification of spatial regions within the brain, which were significantly different on the second examination compared with the first examination of the same patients. There were 4 potential patterns for radiologist readings of these 2 examinations: normal on both, abnormal on both, normal then abnormal, or abnormal becoming normal. Given the timing of the MR examinations relative to therapy, it was not anticipated that the results of any patient examination would be abnormal after induction and would become normal after the more intensive consolidation phase of therapy. Indeed, none of the 197 subjects exhibited this pattern. As would also be expected, most of the patients’ results were read as normal on both examinations (140 [71%] of 197). Of the remaining patients, 31 (16%) of 197 exhibited T2 hyperintensities on the second examination only and 26 (13%) of 197 exhibited changes on both examinations.

Figure 1 demonstrates the voxels that were significantly different at a level of *P* < .00002 with a cluster threshold of 3000 voxels. The *F*-values displayed in the colored clusters have a maximal value of 60. Two large contiguous clusters of approximately the same size and spatial extent were identified (*n* = 19,329 voxels on the left and *n* = 18,689 voxels on the right). These highly significant clusters were most prominent in the frontal lobes and extended superiorly over the ventricles into the posterior parietal lobes. These regions seemed to involve predominantly the anterior, superior, and posterior corona radiata and superior longitudinal fasciculus fiber tracts.

**Two-Group Analysis**

Because the changes during consolidation therapy demonstrated in the first analysis could vary with intensity of therapy and potentially age at treatment, a second 2-group analysis, stratified by treatment risk-group, was performed. Age at the time of treatment was not included in the analysis because it correlated highly with treatment risk-group (*R* = 0.405; *P* < .001). The LR patients were an average age of 5.1 ± 3.1 years at diagnosis, whereas the SHR patients had an average age of 8.6 ± 4.8 years at diagnosis. There was no substantial difference in the proportion of examinations read by the radiologist as exhibiting T2 hyperintensities between the 2 groups, with 28% of LR subjects and 30% of SHR subjects read as abnormal.

Figure 2 demonstrates the voxels that were significantly different at a level of *P* < .0001 with a cluster threshold of 1000 voxels. The *F*-values displayed in the colored clusters have a maximal value of 8, which is substantially lower than the 1-group analysis as expected. Three large clusters were identified, with 2 on the left corresponding to approximately the same size and spatial extent as the 1 on the right (*n* = 16,302 voxels on the anterior left, *n* = 3893 voxels on the posterior left, and *n* = 25,023 voxels on the right). As with the 1-group analysis, these significant clusters were most prominent in the frontal lobes and extended superiorly over the ventricles into the posterior parietal lobes involving predominantly the corona radiata and superior longitudinal fasciculus fiber tracts.

---

Fig 1. VBA results from the 1-group analysis of all patients (*n* = 197) regardless of treatment risk group are displayed on the average of the normalized T2-weighted images from the first examination for visualization of the spatial localization and extent of the regions.
Representative Patient Examination

A representative patient examination is shown in Fig 3. This patient was 3.4 years of age at diagnosis and was treated on the LR group of the protocol. Results of the first MR imaging examination were normal, and the second examination exhibited extensive T2 hyperintensities. Representative sections from the normalized T2-weighted second examination are shown in the figure, along with the corresponding clusters identified by the 2-group VBA at the same spatial locations. It can be readily appreciated that almost all of the T2 hyperintensities seen in this patient are within the locations identified by the VBA. Some areas of hyperintensities for this individual patient examination are larger than the clusters from the VBA. The result of the VBA is a map that identifies the most significant regions exhibiting T2 hyperintensity changes during consolidation therapy and is based on population differences that will not map exactly to every individual patient. However, individual examinations with T2 hyperintensities should exhibit a high degree of spatial overlap with the regions identified by the VBA.

Characteristics of Individual T2 Signal Intensity within the VBA Clusters

A series of analyses were conducted to characterize the individual changes in T2 signal intensity within the VBA identified clusters for symmetry and changes within and between examinations as functions of age at examination. The 2 clusters identified in the 1-group analysis were similar in size and shape, with less than 3% difference in size. Analysis of the average signal intensity from each of these clusters also proved to be highly symmetric as shown in Fig 4. A linear regression analysis of each examination with a zero intercept demonstrated unity slopes with correlation coefficients greater than 0.98. On the basis of this high degree of symmetry in signal intensity, we combined data from the 2 clusters for additional analysis as a function of age at examination.

Figure 5 demonstrates the nonlinear distribution of average T2 signal intensity from the VBA clusters for both examinations on all 197 subjects plotted against age at examination. Data from each examination were then assessed with a logistic regression to demonstrate the longitudinal change in T2 signal
intensity. The 2 regression curves diverge with increasing age at examination so that the increase in signal intensity on the second examination appears greater for older subjects. To more clearly illustrate this effect, Fig 6 was generated on the basis of differences in signal intensity calculated from the 2 logistic regressions evaluated at ages ranging from 2 to 18 years. It was evident from the plot that the change in T2 signal intensity between examinations was twice as great for the 18-year-old compared with the 2-year-old. Furthermore, this effect of age at examination on T2 signal intensity change during therapy was nonlinear.

Discussion
This study identified 2 highly significant bilateral clusters of T2 signal intensity change in WM among patients treated for ALL. The regions were symmetric in size, shape, and average signal intensity. These clusters were most prominent in the frontal lobes and extended superiorly over the ventricles into the posterior parietal lobes involving predominantly the anterior, superior, and posterior corona radiata and superior longitudinal fasciculus fiber tracts. Changes in average signal intensity from these regions both within and between examinations were nonlinear functions of age at examination. T2 signal intensity was greater on the second examination, and the difference between the first and second examination was greater for older subjects. Similar regions were also identified on a 2-group analysis, which controlled for treatment risk-group.

Neuroimaging studies have consistently reported morphologic changes in the cerebrum of children treated for ALL without radiation therapy.3,14-18 Most of these studies have focused on WM changes, particularly in the frontal lobes.3,18 Most recently, a whole brain voxel-based morphometry study of ALL survivors treated without irradiation was compared with healthy control subjects.19 The study found 2 distinct clusters of reduced WM volume in the right frontal lobe. The region identified in the right middle frontal gyrus of these long-term survivors is encompassed within the spatial regions identified in the current VBA of T2-hyperintensity during treatment.

A recent review by Reddick et al5 included a VBA analysis of patients during therapy for ALL stratified by age at diagnosis but did not control for treatment group (intensity of therapy), which correlated highly with age at diagnosis. In the current longitudinal study, all 197 subjects are used in the VBA, and treatment group is included as a covariate, thus increasing the power of the study and resulting in increased statistical significance as demonstrated by the P values achieved. This improved approach also has the advantage of resulting in a single outcome (region of interest) across all subjects rather than 2 results depending on the age of the patient. Another advantage
of the current study is the use of the high-resolution ICBM T2 50 atlas (1-mm isotropic) as opposed to a single-subject examination with lower-resolution 4-mm-thick 2D imaging. The higher resolution target facilitated improved spatial normalization of the patient examinations, resulting in improved sensitivity to longitudinal WM changes. The deterministic atlas space also facilitated the identification of the corona radiata and superior longitudinal fasciculus as fiber tracts involved in the WM abnormality.

Changes in the frontal WM tracts identified by the current VBA could be the cause of subsequent cognitive difficulties in this cohort. Additional changes in the posterior WM tracts may indicate a difference in susceptibility of WM with varying degrees of myelination. Maturation, growth, and organization of regional WM, specifically that in the frontal-parietal regions, play an important role in cognitive development.20 ALL survivors exhibit a consistent pattern of deficits in attention/executive functions, memory (particularly nonverbal), and processing speed,21-25 which are often associated with frontal lobe domains17,26 and are specifically supported by the fiber tracts implicated in this study.

While evaluating the potential impact of these WM changes, one must consider the potentially confounding effect of age at examination on the interpretation of the results. Figures 5 and 6 demonstrate an increasing difference in T2 signal intensity between the examinations for older subjects. Because older ages are associated with high-risk leukemia and, hence, more intensive therapy such as higher doses of methotrexate, we cannot definitively determine if the increasing difference is because of age or intensity of therapy. Although older age cannot be excluded as a potential confounding variable, most studies of neurotoxicity have determined that younger age at therapy is more often associated with increased risk for damage. Furthermore, increased intensity of CNS-directed therapy is often associated with more imaging and neurocognitive difficulties in survivors.

Another potential limitation of this study was the use of the ICBM T2 atlas, which is based on imaging of young adults. Our patient cohort had a median age of only 5.3 years and included children as young as 1 year. The T2 signal intensity of WM across the full age range included in this study changed by almost 25%. A potential of error is associated with the spatial normalization of young undeveloped brains to an adult atlas. The accuracy of the FFD spatial normalization for pediatric cancer patients was established by Zhang et al9 as an average 2-mm error over all predefined manual landmarks. The magnitude of this error is well below the full width half maximum 5-mm Gaussian kernel used to smooth the examinations. To assess the performance of the spatial normalization, each examination was verified visually for gross distortions or misalignments. Although this does not eliminate the possibility of more subtle misalignments of cortical structures, the alignments of deep WM regions were less likely to be affected. This assessment is consistent with the clarity of the average fractional anisotropy (FA) image produced across all subjects.

Methotrexate has been the primary focus of many studies of neurotoxicity and adverse neurocognitive outcomes.27-29 The dose, cumulative exposure, and infusion rate of methotrexate have each been associated with the degree of neurocognitive deficit.27,30 Methotrexate, however, is not the only drug used in the treatment of ALL that can adversely affect the mor-
phologic features of the developing brain.34 For example, asparaginase may result in cerebrovascular disease, especially venous sinus thrombosis. Corticosteroids such as dexamethasone may also have a pronounced effect on cerebral morphologic features.

Although this prospective study focused on the conventional T2-weighted MR signal intensity changes, another more specific approach to characterize microstructural changes in WM during therapy is diffusion tensor imaging, which quantifies the diffusion of water in tissue.32 Clinical studies of childhood ALL and brain tumors have demonstrated that WM FA was reduced following treatment, and that it is more sensitive than conventional T2-weighted MR imaging in the detection of WM injury.33,34 The use of diffusion tensor imaging in asymptomatic patients during or after treatment of ALL is a more recent development. The association of WM FA and IQ in 18 ALL survivors and age-matched healthy control subjects has been prospectively tested.34 The FA in the WM of ALL survivors was often lower than that in the age-matched control subjects, and the differential between the FA in patients relative to that in control subjects was directly proportional to the full-scale, performance, and verbal IQ scores of the survivors. An ongoing longitudinal study of diffusion tensor imaging as a clinically useful approach for the assessment of treatment-related neurotoxicity is being investigated in regions identified by the current VBA analysis. An example of conventional and diffusion tensor imaging from a young patient during therapy for ALL is demonstrated in Fig 7. This imaging is from a time point in therapy that is approximately the same as the second time point in the current study. The region of hyperintensity is again noted in the corona radiata and displays decreased FA, only slight elevation of axial diffusivity, and markedly increased radial diffusivity. This pattern is consistent with a transient demyelination and needs to be investigated prospectively.

Conclusions

In conclusion, this study identified 2 highly significant bilateral clusters of T2 signal intensity change in WM among patients during treatment of ALL with use of both 1-group and 2-group analyses controlling for treatment risk-group. The regions were symmetric in size, shape, and average signal intensity involving predominantly the anterior, superior, and posterior corona radiata and superior longitudinal fasciculus fiber tracts. Changes in average T2-weighted signal intensity from these regions both within and between examinations were nonlinear functions of age at examination, and the difference between the examinations was greater for older subjects who received more intense therapy. Whole-brain VBA identified specific WM tracts that seem to be at increased risk for the development of T2-weighted hyperintensities during therapy for childhood ALL and may be the cause of subsequent cognitive difficulties in survivors. Additional studies will be needed to more fully test this relationship and should include diffusion tensor imaging.

Acknowledgment

We thank Chiquila Jones for her efforts in processing and analysis of the MR imaging examinations.

References


