

A Functional Magnetic Resonance Imaging Study of Language Processing and Its Cognitive Correlates in Prematurely Born Children

Bradley S. Peterson, MD*; Betty Vohr, MD‡; Michael J. Kane, BS*; Douglas H. Whalen, PhD§; Karen C. Schneider, MPH‡‡; Karol H. Katz, MS¶; Heping Zhang, PhD¶; Charles C. Duncan, MD**; Robert Makuch, PhD¶; John C. Gore, PhD||; and Laura R. Ment, MD‡‡#

ABSTRACT. *Objective.* Abnormalities in brain structure, cognition, and behavior have been described in children born prematurely. However, no direct in vivo evidence has yet demonstrated abnormal neural processing in these children. Our aim was to compare brain activity associated with phonologic and semantic processing of language between term and preterm children using functional magnetic resonance imaging (fMRI).

Methods. fMRI scans were acquired during a passive language comprehension task in 26 preterm children at 8 years of age and in 13 term community control children who were comparable in age, sex, maternal education, and minority status. IQ was assessed using a standard measure of intelligence.

Results. The pattern of brain activity identified in a semantic processing task in the preterm children closely resembled the pattern of brain activity identified in a phonologic processing task in term controls. The greater this resemblance in the preterm children, the lower their verbal comprehension IQ scores and the poorer their language comprehension during the scanning task.

Conclusions. Preterm children with the poorest language comprehension seemed not to fully engage normal semantic processing pathways in a language comprehension task. These children instead engaged pathways that normal term children used to process meaningless phonologic sounds. Aberrant processing of semantic content in these preterm children may account in part for their lower verbal IQ scores. *Pediatrics* 2002;110:1153–1162; *functional MRI, children, prematurity, cognition, outcome.*

ABBREVIATIONS. fMRI, functional magnetic resonance imaging; IVH, intraventricular hemorrhage; ADHD, attention-deficit/hyperactivity disorder; AC-PC, anterior commissure-posterior commissure; ROI, region of interest; BA, Brodmann's area; CI, confidence interval.

From the *Columbia College of Physicians and Surgeons and the New York State Psychiatric Institute, New York, New York; ‡the Department of Pediatrics, Brown University School of Medicine, Providence, Rhode Island; §Haskins Laboratories, New Haven, Connecticut; the Departments of ||Diagnostic Imaging, ¶Epidemiology and Public Health, #Neurology, **Neurosurgery, and ‡‡Pediatrics, Yale University School of Medicine, New Haven, Connecticut.

Received for publication Dec 5, 2001; accepted Jun 12, 2002.

Reprint requests to (B.S.P.) Columbia College of Physicians and Surgeons and New York State Psychiatric Institute, Unit 74, 1051 Riverside Dr, New York, NY, 10032. E-mail: petersob@child.cpmc.columbia.edu
 PEDIATRICS (ISSN 0031 4005). Copyright © 2002 by the American Academy of Pediatrics.

Children born prematurely are known to be at risk for cognitive and behavioral problems, including lower IQ, learning disorder, thought disturbance, and excessive distractibility and hyperactivity.^{1–6} Anatomic abnormalities in brain development have been associated with preterm birth^{7–16} and are especially prominent in pre-motor, sensorimotor, and parietooccipital cortices.¹⁶ The magnitude of these abnormalities at 8 years has been reported to correlate with IQ and gestational age at birth.¹⁶ Preterm birth therefore predisposes children to structural disturbances in brain development, which may in turn contribute to poorer cognitive outcomes in middle childhood. Thus far, however, no direct in vivo studies have demonstrated that brain functioning is abnormal in children born prematurely, nor has it been shown whether or how functional abnormalities in the brain are related to cognitive outcome. We were particularly interested in using functional magnetic resonance imaging (fMRI) to study the neural basis of language functioning in preterm children because language deficits have important and far-reaching influences on academic and social functioning throughout childhood development. fMRI has been used extensively to study language processing in adults, and it is being increasingly applied to the study of language functioning in children.

Language is not a unitary cognitive process but one comprising multiple information-processing components that include prosodic, phonologic, syntactic, lexical, semantic, and expressive language functions. fMRI is in principle able to identify brain activity associated with particular cognitive functions, such as each of these linguistic subprocesses, by comparing images acquired during performance of a primary task of interest with images acquired during a second baseline task. The baseline task ideally differs from the primary task of interest by a single (or more realistically a few) cognitive processing elements. When these cognitive features of the tasks are tightly controlled, the brain activity associated with the control task is said to subtract out from brain activity associated with the primary task of interest, in theory leaving only the brain activity associated uniquely with the cognitive processing elements in which the tasks differ.

fMRI subtractions typically are constructed such that each level within a hierarchy of tasks serves as

the control for the task above it.¹⁷ For example, in studies of language processing, a hypothetical hierarchy of tasks might be: 1) gaze fixation, 2) passive viewing of nonsense written symbols, 3) passive viewing of strings of letters from the alphabet (but arranged so as not to be real words), 4) viewing of letter strings that represent real words that are nouns, and 5) viewing of the same words while subjects generate verbs associated with those nouns. The hierarchy of subtractions that corresponds with this hierarchy of tasks would in principle yield language subcomponents akin to: 1) orthographic processing (task 2 – task 1), 2) phonologic processing (task 3 – task 2), 3) semantic processing (task 4 – task 3), and 4) word generation, or expressive language functioning (task 5 – task 4).

No fMRI study can assess all these language functions in a single scanning session because of the prohibitive time needed to obtain imaging data in all the task conditions. These practical limitations are even more formidable when scanning children, especially cognitively and behaviorally disturbed children, who are unable to comply with complicated tasks in long scanning sessions. In this fMRI study of language processing in term and preterm children, we therefore focused on studying 2 broad aspects of language functioning: phonologic processing (the encoding and processing of phonemes, the elemental sounds of speech) and semantic processing (comprehending the cognitive meaning conveyed by speech). We hypothesized that regional brain activity in these language-based tasks would differ across groups consisting of 26 preterm and 13 term children. We also hypothesized that measures of functional activity during a task entailing semantic comprehension would correlate with verbal comprehension IQ scores in the preterm cohort. The results suggest that the neural processing of semantic material in prematurely born children is seriously disturbed and that the more disturbed this neural processing, the lower the verbal IQ scores are in these children.

METHODS

This study was performed at the Yale University School of Medicine in New Haven, Connecticut, and at Brown University School of Medicine in Providence, Rhode Island. Each site's institutional review board approved the procedures. Children and their parents provided written consent and were reimbursed for their time and expenses while participating in the study. MRI scans were obtained at Yale.

Subjects and Their Assessment

The preterm cohort consisted of children enrolled in the follow-up component of a multisite trial of indomethacin to prevent intraventricular hemorrhage (IVH).^{18,19} The first 26 children sequentially enrolled in the prevention study as neonates were recruited for the imaging study when they reached 8 years corrected age. We have previously shown that these children were representative of the entire cohort from which they were selected.¹⁶ The term children, aged 7 to 9 years, were recruited from the local community and group-matched with the preterm group by age, sex, maternal education, and minority status.^{16,20} Four term adult subjects were also studied to assess how comparable the task-related brain activity in term children was with the activity in healthy adults. The assessments of neonatal health status and neurodevelopmental outcome have been described.^{16,18} Blinded assessment of intelligence was performed at 96 months' corrected age using the Wechsler Intelligence Scale for Children-III.²¹ The

severity of attention-deficit/hyperactivity disorder (ADHD) symptoms was assessed using a standardized rating scale.²²

fMRI Subtraction Paradigms

The behavioral tasks used in the fMRI subtraction paradigms were selected to help identify brain regions involved in the phonologic or semantic processing of spoken language. Given the wide range of IQs in these children, their young ages, and the common emotional and behavioral difficulties in preterm children—all of which would make it difficult for the preterm children to comply with an fMRI protocol—it was imperative that the task be enjoyable and appropriate to the range of mental ages for the participating children. It also had to place few behavioral demands on the children to help minimize motion artifact and to reduce variability in brain activations caused by cognitive confusion and poor compliance with task instructions. We therefore selected a passive auditory listening task in which the children listened to 3 varying stimulus presentations of a pleasant children's story.

Stimulus 1

This was an audiotape of a young woman reading in a pleasant voice *The Ugly Duckling*.²³ The story was presented in its entirety during the scanning session, although it was broken into 6 segments, each 35 seconds long. Experts in developmental linguistics confirmed that the vocabulary demands and semantic content were age appropriate for this study population.

Stimulus 2

This was the same children's story but with all the phonemes of the story randomized in time. The story was read by the same woman and with the same prosody as the original (nonrandomized) story. Reading this proficiently took several days' practice with professional audio technicians before audiotaping. This randomized version of the story was designed to contain the same acoustic spectral frequencies, phonemes, prosody, and duration as the original story. Phonemic randomization was intended to destroy the linguistic structure that made semantic comprehension possible. Comparing images acquired while the children listened to the original story (Stimulus 1) with those acquired while they listened to the randomized story (Stimulus 2) therefore should largely identify the differential brain activity needed to process the semantic components of the original story.

Stimulus 3

Here the randomized story was low-pass filtered (ie, high-frequency components of speech sounds were removed) so that its phonemes could not be discerned. This version of the story therefore had the same prosody and duration as the first 2 versions, but it was virtually devoid of phonemic content. Comparing images acquired as the children listened to this low-pass filtered story (Stimulus 3) with images acquired as they listened to the phonemically randomized story (Stimulus 2) therefore should largely identify the differential brain activity associated with processing of the phonologic (but not semantic) components of the story.

Low-pass filtering of the speech stimuli was accomplished using hardware manufactured by Frequency Devices Inc (Haverhill, MA). The cutoff frequencies for the low-pass filter were tunable, and they were implemented using an 8-pole Butterworth filter. The Butterworth transfer function provided a maximally flat amplitude response in the pass band. Attenuation was 3 dB at the cutoff frequency, and it rolled off at 48 dB per octave beyond that. The filter's resolution depended on the frequencies of the acoustic stimuli (with a ratio of 1:299), so that the error was ~1 Hz for the acoustic frequencies used here.

Stimulus Presentation

The audiotaped stimuli were delivered through headphones. Portions of the 3 versions of the story were presented sequentially to the children in an alternating sequence. Each was presented as 2 separate segments (each 35 seconds long) in each of 3 experimental runs (each run 3 minutes, 30 seconds long). The segments were presented sequentially within each story throughout the scanning session, but the order of presentation of the 3 types of

stimuli was pseudorandomized within runs and counterbalanced across runs.

The children were told that they would be listening to a story through the headphones. They were instructed to listen to the story closely and to try to understand and remember what was being said. They were also told, without greater specification, that they would not be able to understand some portions of what they would hear. To assess their comprehension of the intelligible portions of the story, they were asked 10 multiple-choice questions about its content after the scan.

Image Acquisition

Head positioning in the magnet was standardized using the canthomeatal line. A T1-weighted sagittal localizing scan was used to position the axial images. In all subjects 10 axial slices were acquired to correspond with 10 axial sections of the Talairach coordinate system²⁴ oriented parallel to the anterior commissure-posterior commissure (AC-PC) line. The slices were positioned with 2 slices below, 7 slices above, and 1 slice containing the AC-PC line. Slice thickness was a constant 7 mm, and the skip between slices varied between 0.5 and 2 mm to maintain a strict correspondence with the Talairach coordinate system.

Images were acquired on a GE Signa 1.5 Tesla scanner (Milwaukee, WI) equipped with echoplanar imaging hardware (Advanced NMR, Inc, Wilmington, MA). The functional images were obtained with a gradient echo, echo planar imaging pulse sequence having a repetition time = 2060 msec, echo time = 45 msec, flip angle = 60°, 1 excitation per image, 20- × 40-cm field of view, and 64 × 128 matrix, providing a 3.1- × 3.1-mm in-plane resolution. During each run, 102 echoplanar images were acquired in each slice, providing 306 images per experiment and 102 images for each stimulus type. High-resolution, T1-weighted anatomic scans for volumetric measurements were acquired using a sagittal spoiled gradient recall sequence (repetition time = 24 msec, echo time = 5 msec, 45° flip, frequency encoding superior/inferior, no wrap, 256 × 192 matrix, field of view = 30 cm, 2 excitations, slice thickness = 1.2 mm, 124 contiguous slices).

Image Processing

Preprocessing

Studies were excluded if visual inspection indicated that the subject moved >0.5 pixels in any direction or if ghosting artifacts significantly degraded the images. SPM99 was used to correct subject motion and to remove from each pixel's time course the correlations pertaining to first- and second-order motion in the *x*, *y*, and *z* directions.²⁵ Drift of baseline image intensity was removed using an eighth-order high-pass Butterworth filter with a frequency cutoff equal to three-fourths of the task frequency. The time series were filtered once forward and once backward to ensure no change in phase of the signal in relation to the phase of the task. Low-intensity pixels outside the brain were removed, and the images were spatially smoothed using a Gaussian filter with a full width at half maximum of 6.3 mm.

Pixel-wise Calculation of Mean fMRI Signal Differences Across Conditions

The average percentage difference in fMRI signal acquired during the presentation of differing auditory stimuli was calculated at each pixel in each child's scan. This served as an index of the differential brain activity while listening to the randomized compared with the low pass-filtered stories (ie, signal change associated with phonetic processing) or the real compared with the randomized stories (ie, signal change associated with semantic processing). We chose to use these percentage signal changes in statistical analyses because they are less prone to systematic influence by subject motion than are *t* statistics; the latter can be systematically reduced in patient groups, who tend to move more than control groups, thus producing larger signal variances.

Average Signal Differences in Regions of Interest

A region-of-interest (ROI) approach was used in testing differences in activation across groups. The more common voxel-based statistical comparison of 2 groups requires that the average brain anatomies for the 2 groups are similar and will overlay one another with sufficient precision. This requirement clearly was not

met in this population, however, because the groups were previously shown to have large differences in regional brain volumes.¹⁶

A priori ROIs were defined using standard stereotactic coordinates,²⁴ with the exception of 1 thalamic and 3 basal ganglia (caudate, putamen, and globus pallidus) ROIs, which were hand circumscribed on the T1-weighted axial anatomic images. Signal changes were then averaged across all pixels having an absolute value of signal change >0.05% within each stereotactically defined ROI (average signal change in most fMRI experiments is only 3% to 4%). This average across pixels served as a measure of the differential regional signal change associated with the behavioral tasks. These regionwise mean signal differences were then entered as dependent measures in formal statistical tests of the a priori hypotheses.

Regional Volumes

The methods for defining regional brain volumes are described in detail elsewhere. Briefly, however, definition of cortical regions was performed with 3 coronal planes (1 tangent to the genu of the corpus callosum, 1 each through the AC and PC) and 1 axial plane (containing both the AC and PC) that divided the cerebrum into stereotactically defined sectors.^{16,26} Basal ganglia, amygdala, and hippocampus regions were defined using manual tracing techniques.^{27,28}

Statistical Analyses

Hypothesis Testing

All statistical procedures were performed in SAS v.8.0 (SAS Institute Inc., Cary NC), and all *P* values were of the 2-sided type. A priori hypotheses were tested in a mixed models analysis (PROC MIXED) using repeated measures (average regional signal differences) over the brain. Average regional signal differences were entered as dependent variables into a single statistical model that included 2 tasks (semantic and phonetic processing), 2 hemispheres (left and right), and 29 regions as within-subject factors, and group (preterm versus term) as a between-subjects factor, with age and sex as covariates. The repeated-measures analysis accounted for correlations between regions within subjects and for correlations between tasks across subjects. Use of the average signal differences for the semantic and phonetic processing tasks as the dependent measures was preferred over the use of mean signal differences in the 3 stimulus conditions because it directly compared activations in the 2 maps of signal differences that are standard in representing differential brain activity during cognitive tasks. Moreover, use of difference measures in repeated-measures analyses is standard in the statistical analysis of 2 or more active conditions that are compared with a single baseline.²⁹ Included in the model were each of the 2- and 3-way interactions for the region, task, and group terms. The model therefore was hierarchically well-formulated.³⁰ We also considered the interactions of hemisphere with these terms, but they did not approach significance and therefore were not included in the final model.

Our first a priori hypothesis was that regional brain activity across these language-based tasks would differ between the term and preterm children. The term of the statistical model that tested this hypothesis was the 3-way interaction of brain region, task, and group (the region-by-task-by-group interaction). To identify the component terms that contributed most to the significance of this 3-way interaction, we examined the parameter estimates and associated *P* values of the component terms in an analysis of fixed effects for the final mixed model. Least-squares means and standard errors were calculated in the mixed models and plotted to assist in the interpretation of significant interactions.

Our second a priori hypothesis was that brain activity during semantic processing would correlate with verbal comprehension IQ scores in the preterm cohort. This was tested by calculating the *P* value for the correlation of verbal IQ with the average percentage change in fMRI signal associated with semantic processing in Wernicke's area of the preterm children. Signal change in Wernicke's area during the semantic processing task was selected for use in this correlation analysis because Wernicke's area has been shown to activate robustly during language comprehension tasks.³¹ The signal change in Wernicke's area was calculated by averaging signal change across all pixels in the ROIs corresponding as closely as possible to Brodmann's areas (BAs) 22, 37, and 39,

located in the posterior portion of the superior temporal, angular, and marginal gyri.³¹

Assessment of Possible Confounds

We also included in the initial models maternal education and previous treatment with indomethacin. Neither of these variables was associated significantly with brain activation, however, and they had negligible effects on the parameter estimates. Therefore, they were not included in the final models for hypothesis testing.

Concordance of Activation Maps

The similarity in activation profiles for the term and preterm children across phonologic and semantic processing was assessed in a quantified analysis of concordance of the group activation maps. Each pixel in each group activation map was assigned a value of -1, 0, or +1 based on its being, on average, a deactivation, nonactivation, or activation, respectively, in the group maps at the above-specified thresholds for percentage signal change and clustering. Weighted kappa coefficients and 95% confidence intervals (CIs) were then calculated to assess the concordance of the activation profiles across all pixels of all slices for each possible pair of group activation maps.³² The possible range of this statistic is from -1 to +1, with zero indicating a chance level of concordance.

Exploratory Analyses

The associations of verbal comprehension and verbal IQ scores with regional activations were assessed in all predefined ROIs in the preterm group using Pearson correlation coefficients. The associations of regional activations with regional brain volumes were similarly assessed.

RESULTS

Subjects

The scans of 5 subjects (3 preterm, 2 term) were excluded from further analyses because of motion or ghosting artifact, leaving 23 preterm and 11 term children in the final sample. Neonatal characteristics of the preterm group are shown in Table 1. Demographic, neurologic, cognitive, and behavioral characteristics at 8 years of age are shown in Table 2. The 4 adult subjects consisted of 1 man and 3 women (all right-handed, 36.5 ± 18 years old).

Hypothesis Testing

Statistical modeling detected a significant task-by-group interaction ($F_{1,32} = 11.3$; $P < .002$). The least-squares means for the task-by-group interaction (Fig 1) indicated that, on average across all ROIs, the term

children deactivated (ie, the mean signal decreased relative to the baseline condition) more in the phonetic processing than in the semantic processing task. The preterm children, in contrast, deactivated more in the semantic than in the phonetic processing task.

In addition to these regionally nonspecific effects reflected in the task-by-group interaction, the model also detected a significant task-by-region-by-group interaction ($F_{28,896} = 1.60$; $P < .03$), indicating regionally specific group differences in task-related fMRI signal change, and confirming our first hypothesis. The analyses of fixed effects indicated that the region- and task-specific group differences were most prominent in BAs 9 ($F_{896} = 1.72$; $P = .08$) and 23 ($F_{896} = 1.79$; $P = .07$), the cerebellum ($F_{896} = 2.14$; $P = .03$), and the globus pallidus ($F_{896} = 1.87$; $P = .06$). The least-squares means for these effects are presented in Fig 2.

Group Average Activations Maps

Maps representing the pixelwise group average signal changes for semantic and phonologic processing are presented in Fig 3. The pattern of activation in the term children was similar to that seen in the adults (not shown). Comparison of the maps of the term and preterm children suggested that the term children tended to activate numerous cortical regions during semantic processing, including receptive and expressive language regions. During phonologic processing, the term children tended to deactivate broad expanses of prefrontal cortex (ie, these regions were more active as the children listened to the prosodic, or low-pass filtered, stimulus). The preterm children, in contrast, tended to deactivate during semantic processing in the same regions as those that deactivated for the term children during phonologic processing. The similarity in activation maps for preterm children during the semantic processing task and term children during the phonologic processing task was confirmed quantitatively in the analyses of concordance of activation maps using the weighted kappa statistic.

IQ Correlations

In the preterm cohort only, the correlation of verbal comprehension scores with activity in Wernicke's area during semantic processing was statistically significant ($r = .54$; $P = .01$), confirming our second hypothesis. Post hoc analyses indicated that the correlation of verbal comprehension scores with semantic processing activity in the preterm cohort was not unique to Wernicke's area but that it was also seen in each component of Wernicke's area (BAs 22, 37, and 39), as well as in BAs 6, 8, 21, and in the sensorimotor cortices (Table 3). These correlations were not significant in the term children. In addition, semantic processing activity in each instance was positively correlated with verbal comprehension scores in the preterm children. Scatter plots suggested that these positive associations, combined with the greater range of IQ scores in the preterm cohort (especially those scores that extended well into the subnormal range), was responsible for the overall tendency for signals in the preterm group to decrease more than

TABLE 1. Preterm Neonatal Characteristics

Neonatal Characteristics	Value
Gestational age (wk)	28.8 (1.7), range: 26-33
Birth weight (g)	999.0 (181.3), range: 700-1240
Multiple gestation pregnancies	3
1-Min Apgar median	5, range: 0-9
5-Min Apgar median	7, range: 3-9
Antenatal steroids	3
Prophylactic indomethacin	12
Grade 1-2 IVH (first 6 h)	4
Subsequent grade 1 IVH	3
Periventricular leukomalacia	1
Ventriculomegaly (at term)	2, mild severity
Number of days on ventilator	10.7 (13.7), range: 0-37
Bronchopulmonary dysplasia	12
Chronic lung disease	3
Number of septic episodes	0.48 (0.99), range: 0-4

Neonatal characteristics of the 23 preterm children with usable fMRI scans were obtained prospectively. Standard deviations are in parentheses.

TABLE 2. Follow-up Characteristics

8-Year-Old Characteristics	Preterm	Term	Significance
Sex	14 boys, 9 girls	5 boys, 6 girls	$\chi^2 = 50, df = 1, P = .72$
Age (y)	8.6 (.38)	8.7 (.92)	$t = .66, df = 32, P = .52$
Minority status	7/23	2/11	$\chi^2 = .57, df = 1, P = .45$
Maternal education (y)	13.9 (2.8), range: 10–20	14.2 (2.7), range: 11–20	$t = .79, df = 32, P = .79$
Height (cm)	127.3 (8.7)	132.6 (5.7)	$t = 1.8, df = 32, P = .08$
Full scale IQ	97.3 (15.0), range: 74–126	107.2 (15.4), range: 90–137	$t = 1.8, df = 32, P = .08$
Verbal IQ	99.5 (15.5), range: 73–128	106.6 (14.6), range: 84–132	$t = 1.3, df = 32, P = .21$
Performance IQ	95.3 (14.6), range: 71–121	107.5 (18.4), range: 84–145	$t = 1.9, df = 32, P = .07$
Verbal comprehension IQ	100.6 (17.0), range: 72–134	112.4 (15.5), range: 81–136	$t = 1.8, df = 32, P = .08$
Freedom from distractibility IQ	97.6 (12.5), range: 75–118	103.1 (7.4), range: 93–115	$t = 1.2, df = 32, P = .25$
Handedness	Right	Right	
Head circumference (cm)	51.5 (1.3), range: 49.5–54.5	NA	
Cerebral palsy	0	0	
Seizures necessitating medication	1	0	
ADHD	5	1	
Separation anxiety disorder	6	0	
Simple phobia	5	0	
Learning disability	2	0	
Chronic motor tics	2	0	

NA indicates not available.

SDs are in parentheses. ADHD was newly diagnosed in 1 term control based on the structure diagnostic interview. The minorities in the study were mostly African Americans but also included several Hispanic and Asian Americans. All *P* values are 2-tailed.

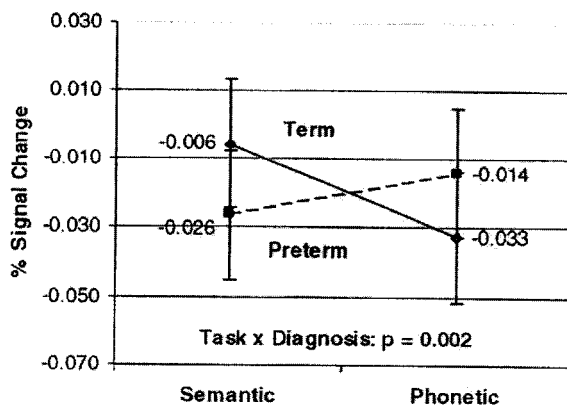


Fig 1. Least-squares means for the task-by-group interaction. The bars represent standard errors. The graph indicates that term children activated more than the preterm children during semantic processing. During phonologic processing, the term children deactivated more than the preterm children. These group differences in task-related activation were independent of brain region.

the signals in term children during the semantic processing task (Fig 4). Correlations of regional activation with the other IQ scales were in the same direction as the correlations of brain activation with verbal comprehension IQ scores, but they were smaller in magnitude. They were strongest for verbal IQ, then for full-scale IQ and freedom from distractibility scores, and weakest for performance IQ scores (data not shown).

Correlations With Comprehension Scores

Term and preterm children did not differ significantly in their comprehension of the story ($63.6 \pm 35.6\%$ vs $61.7 \pm 36.1\%$ correct, respectively). In the preterm group, story comprehension scores (percentage of questions answered correctly) correlated most strongly with verbal IQ ($r = .63; P = .001$) and verbal comprehension IQ ($r = .62; P = .001$), but they also

correlated with full-scale IQ ($r = .45, P = .03$) and freedom from distractibility IQ ($r = .52; P = .01$). Story comprehension scores did not correlate significantly with performance IQ ($r = .16; P = .44$), perceptual IQ ($r = .22; P = .31$), or perceptual processing speed IQ scores ($r = .03; P = .90$). These correlation analyses support the validity of the story comprehension scores obtained in this sample as a measure of cognitive understanding of the story.

Story comprehension scores in the preterm group did not correlate significantly with a single region's activation during phonologic processing, and yet they did correlate significantly with numerous regional activations during semantic processing, including BA 8 ($r = .57; P = .009$), BA 6 ($r = .53; P = .009$), the sensorimotor cortex ($r = .55; P = .007$), BA 7 ($r = .49; P = .02$), BA 11 ($r = .51; P = .01$), BA 39 ($r = .44; P = .04$), and BA 22 ($r = .43; P = .04$). The scatter plots representing these associations were similar to those for the associations of regional brain activity and IQ scores (Fig 4). As was true for the IQ scores, each significant correlation of semantic processing activity with story comprehension scores indicated that higher comprehension scores accompanied greater activations during the semantic processing task.

Concordance of Activation Maps

The weighted kappa for the concordance of preterm semantic and term phonologic group activation maps was calculated to be 0.42 (CI 0.38–0.46), indicating moderate concordance. This was far better than the kappa's measuring concordance for any of the other pairs of maps, which did not differ significantly from one another (although they all did differ significantly from zero, suggesting nonspecific common activation across all tasks): preterm semantic versus term semantic = 0.08 (95% CI: 0.05–0.10), preterm phonologic versus term phonologic = 0.03 (95% CI: 0.002–0.06), preterm phonologic versus

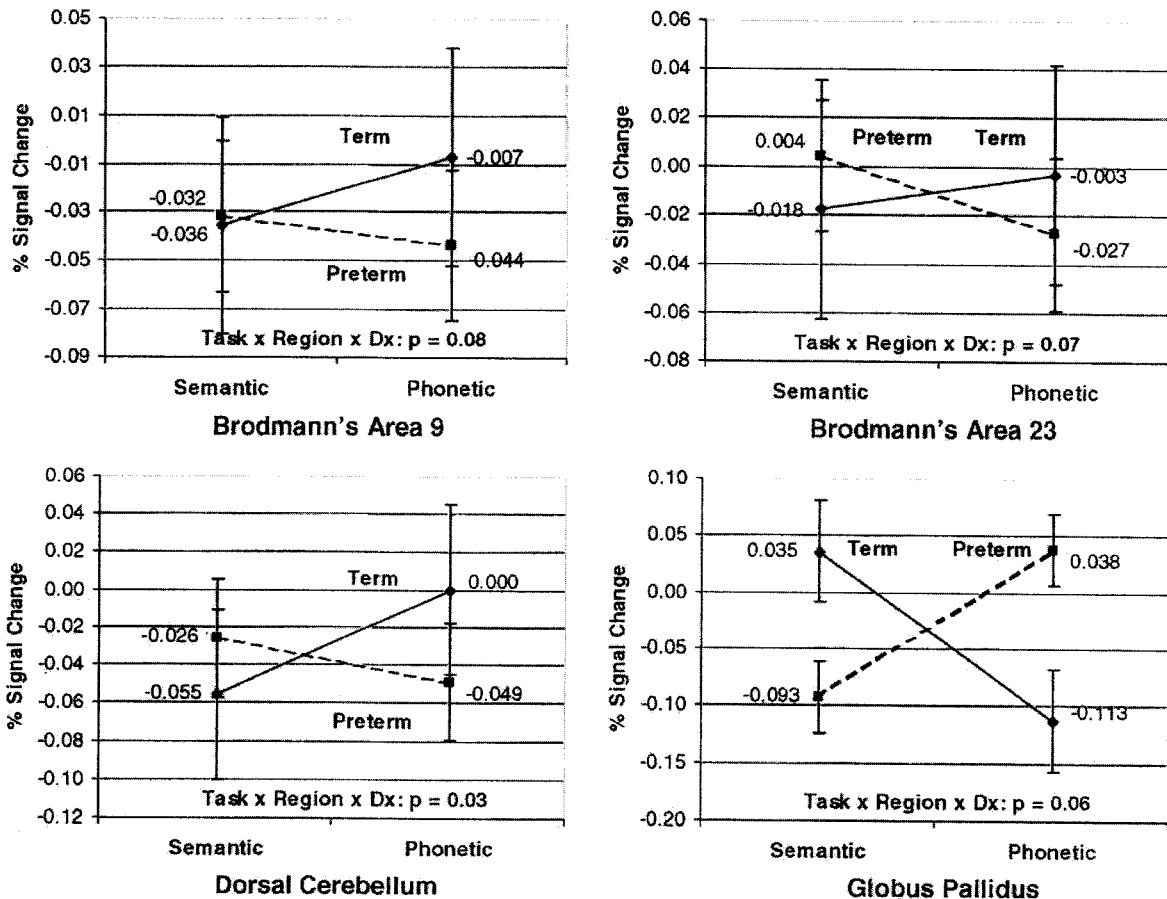


Fig 2. Least-squares means for selected fixed effects. These graphs represent regions in which the fixed effects analyses suggested that they contributed most to the overall significance of the task-by-region-by-group interaction. In general, regional activation during semantic processing in the preterm children was similar to the activation during phonologic processing in the term controls.

term semantic = 0.07 (95% CI: 0.04–0.09), preterm semantic versus preterm phonologic = 0.03 (95% CI: 0.003–0.05), and term semantic versus term phonologic = 0.07 (95% CI: 0.04–0.09).

Exploratory Analyses

Regional activations in either task did not correlate with regional volumes of the cerebral cortex, basal ganglia, amygdala, or hippocampus in the preterm group more often than would be expected by chance. Activations also did not correlate significantly with perinatal variables or neuropsychiatric diagnoses.

DISCUSSION

Analyses of brain activity associated with the phonologic and semantic processing of a children's story demonstrated that 8-year-old term and preterm children differ in their neural processing of language-based tasks. In addition, brain activity during semantic processing correlated significantly with verbal comprehension IQ scores in preterm but not term children. Formal statistical testing of group differences in regional brain activity and the visual representation of average brain activity in each group suggested that the preterm children on average were

processing semantic material using the same neural pathways as those used for phonologic processing in term children. In colloquial terms, preterm children seemed to process meaningful speech in the same way that term children processed strings of meaningless sounds. The more this was true for the preterm children, the poorer was their comprehension of the meaning of the story and the poorer were their verbal subscale and verbal comprehension IQ scores.

During semantic processing, the term children tended to activate both receptive and expressive language areas bilaterally (Fig 3), consistent with previous findings in children and adults during language comprehension tasks.^{33–41} During phonologic processing, in contrast, the term children tended to activate mesial prefrontal cortices and to deactivate both lateral prefrontal and ventral anterior cingulate regions. These deactivations during phonologic processing indicated that these brain regions were more active as the term children listened to the prosodic elements (ie, the low pass-filtered portions) of the children's story compared with brain activity when they listened to its phonemically randomized portions.

In contrast, the preterm children tended to activate

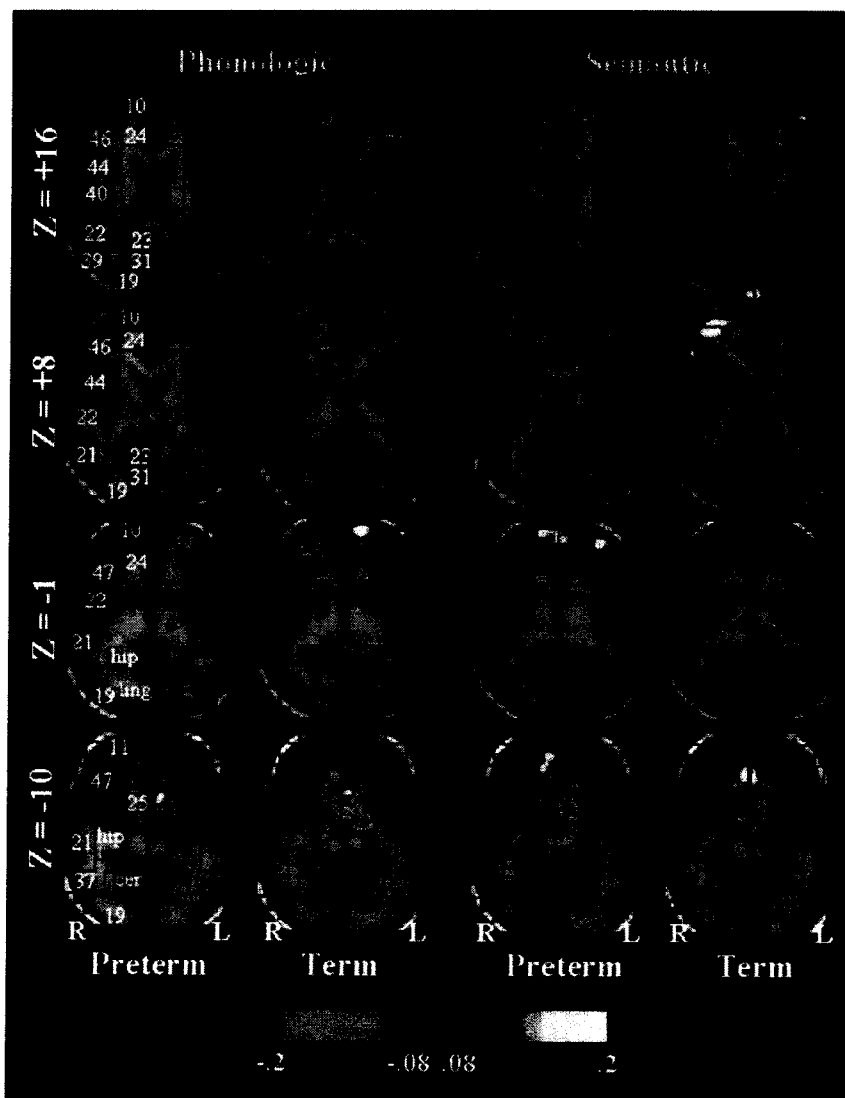


Fig 3. Group activation maps. Shown here are the group activation maps for semantic and phonologic processing for the term and preterm children. Each subject's T1-weighted anatomic images underwent piecewise linear warping to a common bounding box and coregistration to the anterior commissure and posterior commissure.²⁴ The same transformations were applied to the corresponding functional maps, which were then averaged across subjects. The group average maps were thresholded at a mean difference in signal change of 0.08% and then filtered so that only clusters of 25 adjacent pixels that survived the threshold would be depicted. The z values indicate the z coordinates (axial level) of the Talairach coordinate system.²⁴ The color bar depicts colors associated with percentage differences in signal change. The green letters and arrows refer to regional activations that contributed most to the similar appearances of the semantic processing maps in preterm children and the phonologic processing maps in terms controls. a, prefrontal cortex deactivation (BA 10); b, prefrontal cortex deactivation (BAs 10 and 46); c, prefrontal cortex deactivation (BA 47); d, prefrontal cortex activation (BA 10); e, ventral anterior cingulate deactivation (BA 25). Only the 4 slices with the greatest extents of activation are shown because of space limitations.

mesial prefrontal cortex and to deactivate both lateral prefrontal and ventral cingulate cortices during the semantic processing task. This pattern of activity was similar to the pattern of activity exhibited by the term children during processing of phonologic material. These task-related differences in brain activity were detected as a significant task-by-group interaction in formal statistical modeling. This interaction indicated that the overall level of brain activity in the term children was generally greater during semantic than during phonetic processing, whereas it was

similar between the 2 conditions in the preterm children (Fig 1). Furthermore, regional activity during semantic processing was similar to the corresponding regional activity in term children during phonologic processing (Fig 2). The similarity in regional activity in preterm children during the semantic processing task and term children during the phonologic task was evident in the 2 group average activation maps (Fig 3). The similarity of the activation patterns in these 2 maps was confirmed with a rigorous statistical measure of concordance.

TABLE 3. Correlations of Regional fMRI Activations with Verbal Comprehension IQ Scores

Region	Side	BA	r	P Value
Wernicke's	Summed	22,37,39	.54	.01
Sensorimotor	Left	1,2,3	.51	.01
Motor	Right	4	.42	.05
Supplementary motor	Right	6	.54	.009
Supplementary motor	Left	6	.56	.007
Dorsal prefrontal	Right	8	.61	.003
Dorsal prefrontal	Left	8	.48	.02
Middle temporal gyrus	Right	21	.42	.05
Superior temporal gyrus	Right	22	.47	.02
Posterior cingulate	Left	23	.47	.02
Limbic cingulate	Left	37	.53	.01
Angular gyrus	Right	39	.62	.002
Angular gyrus	Left	39	.51	.01

r indicates Pearson's correlation coefficients for regional brain activity correlating with Verbal Comprehension IQ scores within the preterm group only (values were similar if normal controls were included). Activations in Wernicke's area were summed across BAs 22, 37, and 39 bilaterally. Only correlations with *P* values < .05 (uncorrected) are shown. No regional brain activities in the term children correlated significantly with IQ scores.

The aberrant neural processing of semantic material in the preterm children had important cognitive correlates in this study. During the semantic processing task, activity in Wernicke's area and in all of its component Brodmann's areas correlated inversely with verbal comprehension IQ and story comprehension scores. The same was true for other brain regions that tended to deactivate during semantic processing in the preterm children, indicating that the more these children processed the semantic material in ways that term children processed phonologic material, the more impaired was their understanding of the story's meaning. These findings suggest that preterm children who are particularly cognitively impaired may not have made sense of the story because they tended to hear the normal story more like the term children heard the randomized story: as meaningless, phonemic sounds. Moreover, the correlations of brain activity with verbal comprehension IQ scores indicated that this abnormal semantic processing probably was not confined to the scanning session and that it might instead have far-reaching and enduring functional consequences in other contexts. In other words, it may be a trait cognitive deficit in the preterm children.

It should be noted that the unusual pattern of brain activity observed in preterm children during semantic processing may be a feature associated more with their lower IQs than with their preterm birth per se (ie, even term children with lower IQs might demonstrate this pattern of brain activity). The correlations of brain activity with IQ differed between term and preterm children, however, suggesting that the association of brain activity with IQ was unique to the preterm children. Even if these associations ultimately prove not to be unique to the preterm group and are instead observed in other neurodevelopmental disorders, or if they are observed at the lower end of the IQ distribution in the general

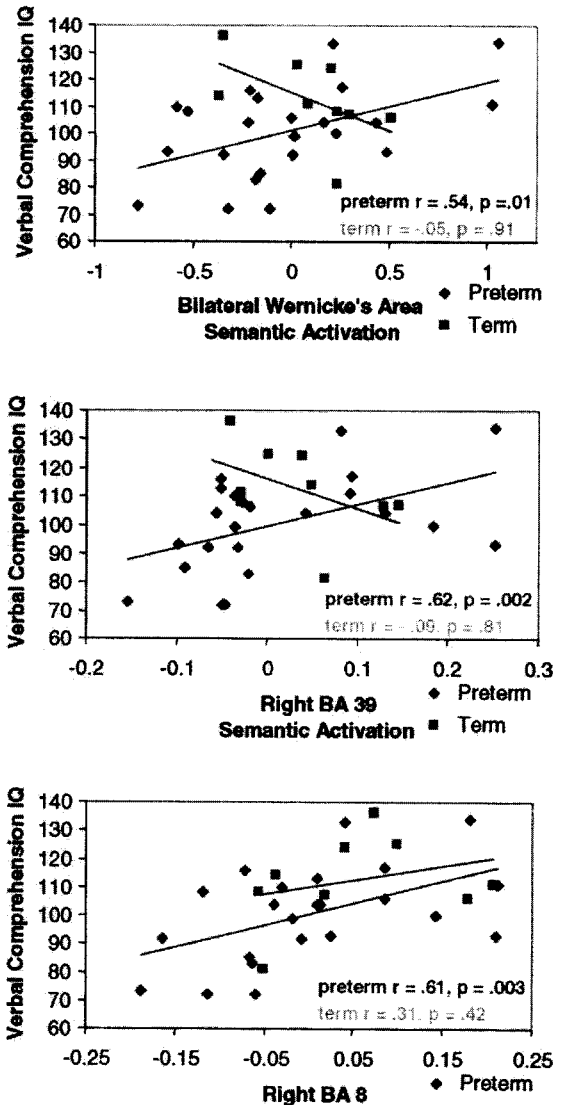


Fig 4. Scatter plots of selected associations of regional semantic activation with verbal comprehension IQ scores. Regression lines are shown for the preterm and term children separately. Greater deactivation in these regions was associated with lower verbal comprehension IQ scores in the preterm children. The lower IQs in the preterm children seemed to be responsible for the tendency of the preterm children as a group to deactivate during semantic processing. The pattern of deactivation during semantic processing in the preterm children was similar to the pattern of deactivation during phonologic processing in the term controls. These findings together suggest that the preterm children with lower verbal comprehension IQ scores were processing semantic material using the same neural systems as term children used when processing phonologic material. The direction of association between brain activation and IQ tended to differ between term and preterm children, particularly in receptive language areas (BA 39 and Wernicke's area), where the group-by-brain-activity interaction was associated with IQ at a trend level of significance ($t > 1.90$; $P < .07$). Values on the x axis are percentage differences in signal as the children listened to the original story compared with values when they listened to the phonemically randomized story.

population, they will still have important implications for our understanding of the neural basis of low IQ.

Conceivably, attentional deficits could have contributed to the apparent tendency of the preterm children to process semantic content as nonsensical sounds. Preterm children simply may not have been attending to the meaning of the story. Several considerations argue against this possibility, however. First, even passive, preattentive processing of words has been shown to activate receptive and expressive language areas in functional imaging studies.^{34,37} Second, the correlations of regional brain activity with verbal comprehension IQ scores in the preterm children were consistently more robust than the correlations with freedom from distractibility IQ scores. Third, neither a categorical diagnosis of ADHD nor ratings of ADHD symptom severity in the preterm children were associated with regional brain activity. Fourth, the story comprehension scores were similar in the term and preterm groups, which would be unlikely if the preterm children were grossly more inattentive than their term counterparts.

Regional brain activity did not correlate with regional brain volumes in the preterm cohort. This may have been caused by a necessarily poor correspondence between the methods that defined regional brain volumes and those that defined the regional brain activities in these children. Alternatively, it could have been caused by a weak or nonlinear association between regional brain volume and neural activity during language processing tasks. Because regional brain volumes and brain activity both correlated with IQ in these children, however,¹⁶ it seems likely that both brain structure and task-specific brain activity contributed unique components to the interindividual differences in IQ scores. This possibility is supported by the relative specificity of the correlations with IQ subscales: brain activity during language processing correlated with verbal but not performance IQ scores, whereas brain volumes correlated with both subscales (and perhaps more consistently with performance IQ scores).¹⁶

Finally, acoustic noise from the scanner may have affected the magnitude of the cerebral activations associated with the auditory tasks used in this study.^{42,43} Nevertheless, the effects of acoustic noise on brain activations in a given task are thought to be stable when the level of acoustic noise is also stable.⁴³ Therefore, because the fMRI scanning pulse sequence was constant across tasks and across subjects, scanner noise is not likely to have affected the differences in cerebral activations across tasks and across groups that we are reporting here.

These findings contribute to the growing evidence that preterm birth is associated with long-term abnormalities in brain development. Those abnormalities now seem to include regionally specific disturbances in brain structure and function. The functional brain disturbances detected here include aberrant semantic processing during a language comprehension task. Patterns of brain activity and measures of comprehension during this task together suggest that preterm children with cognitive impairment tend to process and to hear meaningful language as though it were meaningless.

Structural and functional abnormalities are not present in all preterm children. Nevertheless, when the abnormalities are present, they seem to be associated with deficits in specific cognitive domains. Because several cognitive processes undoubtedly contribute to semantic understanding after phonologic awareness is established (eg, syntactic processing, lexical memory retrieval, and affective and polysensory associations), additional study is needed to determine which of these semantic processing components are uniquely disrupted by preterm birth. Additional studies are also needed to determine whether the patterns of brain activity observed during language processing tasks can be useful diagnostically or in monitoring educational or therapeutic interventions.

ACKNOWLEDGMENTS

This work was supported in part by grants NS 27116 and NS 35476 (NINDS); RR 06022 (National Center of Research Resources); MH30929, MH01232, and MH59139 (NIMH); HD-01994 (Child Development); DC-00493 (National Institute on Deafness and Other Communication Disorders); and the Suzanne Crosby Murphy Endowment at Columbia University.

We thank Drs Giovanna Spinella, Walter Allan, and William Oh for their scientific expertise; Marjorie Ainley, Lisa Perry, and Terry Leach for neurodevelopmental testing; and Hedy Sarofin, Terry Hicky, and Cheryl Lacadie for their technical assistance.

REFERENCES

1. Wolke D. Annotation: supporting the development of low birthweight infants. *J Child Psychol Psychiatry*. 1991;32:723-741
2. McCormick M, Workman-Daniels K, Brooks-Gunn J. The behavioral and emotional well-being of school-age children with different birth weights. *Pediatrics*. 1996;97:18-25
3. Hack M, Friedman H, Fanaroff AA. Outcomes of extremely low birth weight infants. *Pediatrics*. 1996;98:931-937
4. Sykes DH, Hoy EA, Bill JM, McClure BG, Halliday HL, Reid MM. Behavioral adjustment in school of very low birthweight children. *J Child Psychol Psychiatry*. 1997;38:315-325
5. Whitaker AH, Van Rossem R, Feldman JF, et al. Psychiatric outcomes in low-birth-weight children at age 6 years: relation to neonatal cranial ultrasound abnormalities. *Arch Gen Psychiatry*. 1997;54:847-856
6. Wood NS, Marlow N, Costeloe K, Gibson AT, Wilkinson AR. Neurologic and developmental disability after extremely preterm birth. *N Engl J Med*. 2000;343:378-384
7. Olsén P, Vainionpää L, Pääkkö E, Korkman M, Pyhtinen J, Järvelin MR. Psychological findings in preterm children related to neurologic status and magnetic resonance imaging. *Pediatrics*. 1998;102:329-336
8. Stewart AL, Rifkin L, Amess PN, et al. Brain structure and neurocognitive and behavioral function in adolescents who were born very preterm. *Lancet*. 1999;353:1653-1657
9. Murphy DJ, Squier MV, Hope PL, Sellers S, Johnson A. Clinical associations and time of onset of cerebral white matter damage in very preterm babies. *Arch Dis Child*. 1996;75:F27-F32
10. Olsén P, Pääkkö E, Vainionpää L, Pyhtinen J, Järvelin M. Magnetic resonance imaging of periventricular leukomalacia and its clinical correlation in children. *Ann Neurol*. 1997;41:754-761
11. Hüppi PS, Schuknecht B, Boesch C, et al. Structural and neurobehavioral delay in postnatal brain development of preterm infants. *Pediatr Res*. 1996;39:895-901
12. Hüppi PS, Maier SE, Peled S, et al. Microstructural development of human newborn cerebral white matter assessed in vivo by diffusion tensor magnetic resonance imaging. *Pediatr Res*. 1998;44:584-590
13. de Vries LS, Eken P, Groenendaal F, van Haaster IC, Meiners LC. Correlation between the degree of periventricular leukomalacia diagnosed using cranial ultrasound and MRI later in infancy in children with cerebral palsy. *Neuropediatrics*. 1993;24:263-268
14. Inder TE, Hüppi PS, Warfield S, et al. Periventricular white matter injury in the premature infant is followed by reduced cerebral cortical gray matter volume at term. *Ann Neurol*. 1999;46:755-760
15. Cooke RWI, Abernethy LJ. Cranial magnetic resonance imaging and school performance in very low birth weight infants in adolescence.