Original Research

Automated Ischemic Lesion Detection in a Neonatal Model of Hypoxic Ischemic Injury

Nirmalya Ghosh, PhD,[1] Rebecca Recker, BA,[1] Amul Shah, BS,[1] Bir Bhanu, PhD,[3] Stephen Ashwal, MD,[1] and Andre Obenaus, PhD[1,2,4,5]*

Purpose: To develop and compare an automated detection system for ischemic lesions in a neonatal model of bilateral carotid artery occlusion with hypoxia (BCAO-H) from T2 weighted MRI (T2WI) to the currently used "gold standard" of manual segmentation.

Materials and Methods: Forty-three P10 BCAO-H rat pups and 8 controls underwent T2WI at 1 day and 28 days. A computational imaging method, Hierarchical Region Splitting (HRS), was developed to automatically and rapidly detect and quantify 3D lesion and normal appearing brain matter (NABM) volumes.

Results: HRS quantified lesion and NABM volumes within 15 s in comparison to 3 h for its manual counterpart, with a high correlation for injury ($r^2 = 0.95$; $P = 8.6 \times 10^{-7}$) and NABM ($r^2 = 0.92$; $P = 1.4 \times 10^{-22}$). Average lesion volumes for mild, moderate, and severe injuries were 3.85%, 28.85%, and 52.98% for HRS and 0.51%, 24.22%, and 48.74% for manual detection. Lesion volumes and locations were similar for both methods (sensitivity: 0.82, specificity: 0.86, and similarity: 1.47).

Conclusion: HRS is an accurate, objective, and rapid method to quantify injury evolution in neonatal hypoxic ischemic injury models.

Key Words: global cerebral ischemia; magnetic resonance imaging; automatic segmentation; automatic lesion detection; hierarchical region splitting
J. Magn. Reson. Imaging 2011;33:772–781.
© 2011 Wiley-Liss, Inc.

[1]Department of Pediatrics, Loma Linda University, Loma Linda, California, USA.

[2]Department of Radiology, Loma Linda University, Loma Linda, California, USA.

[3]Center for Research in Intelligent Systems, University of California, Riverside, California, USA.

[4]Department of Radiation Medicine, Loma Linda University, Loma Linda, California, USA.

[5]Department of Biophysics and Bioengineering, Loma Linda University, Loma Linda, California, USA.

*Address reprint requests to: A.O., Non-Invasive Imaging Laboratory, Radiation Medicine Department, Loma Linda University, 11175 Campus Street, CSPA1010, Loma Linda, CA 92354.
E-mail: aobenaus@llu.edu

NEONATAL HYPOXIC ISCHEMIC injury (HII) occurs in 3–5 newborns per 1000 live births and may lead to devastating sequelae including cerebral palsy, epilepsy, and cognitive delay (1) and numerous animal models simulating HII have been developed (2). MRI is increasingly used to noninvasively assess the evolution, pathogenesis, and severity of focal and global HII in human newborns and animal studies (3–5). Simplified semi-quantitative scoring systems to grade injury severity from MRI have been developed for clinical use (6,7) and more recently these scoring systems have been adapted for use in rat pups (8). These scoring systems, although helpful, are subjective and qualitative measures of brain injury and often do not predict injury volumes. Likewise, manual quantification of HII lesions is extremely time-consuming, has low reproducibility between observers (individuals may classify the same MRI region differently), and may be error-prone due to fatigue-related errors (influence of neighboring data may affect the true value of a particular pixel/voxel) (9,10).

With the advent of high resolution and multi-modal medical imaging (11), one critical challenge is to objectively analyze such medical and biological information in as reproducible, quantifiable and accurate manner as possible at or near real-time speed. This would greatly improve serial data acquisition in clinical studies and in animal models and improve lesion volume determination as the injury evolves over time (11). Advances in computer vision and computational analysis can enhance automated extraction and quantification of such data and these approaches can greatly improve our ability to monitor ongoing biological processes (12–15).

In recent years, research has been focused on automated MRI segmentation of anatomical brain structures (gray matter, GM; white matter, WM; and cerebrospinal fluid, CSF) using a variety of methods including fuzzy connectedness (16), constrained Gaussian mixture models (GMM) (17), deformable models with K-means clustering (18), spatially constrained fuzzy kernel clustering (14), graph cuts with tissue prior models (19), and hidden Markov models (HMM) (20). In some cases, the lesion has been modeled as a reject class (21) or outliers of the normal tissue models. For example, voxels not satisfying the

criteria of GMM normal tissue models derived from the Expectation Maximization (EM) algorithm (22) are detected as lesions. MRI based automated lesion detection in other neurological diseases, such as multiple sclerosis (15,23–25), focal cortical dysplasias (26,27) and ischemic stroke (11) has generated great interest. There are no studies that have examined the use of automated detection of neonatal ischemic injury (the focus of this work). In part this is due to the increased difficulty due to the continuous evolution of the ischemic lesion in neonates and the unavailability of reliable anatomical templates of the developing brain. A range of strategies have been applied to measure lesional spatial regularization or congruity including MRI textures (13,26), probabilistic models (28), morphological operations (29), and physical model estimation (30), most of which are computationally intensive and not useable for quick injury estimation.

We have developed an automated computational method, *Hierarchical Region Splitting* (*HRS*) that is computationally efficient yet detects neonatal brain ischemic injury comparable to manual segmentation (and other established methods). Most importantly, HRS provides a generic framework that does not require a brain atlas or probabilistic disease model. We investigated the ability of HRS to detect (in three-dimensions) the HII lesion and normal appearing brain matter (NABM) volumes and compared our computational approach with manual volumetric methods in a rat pup model of bilateral common carotid artery occlusion with hypoxia (BCAO-H) that mimics global HII seen in term newborns (8).

Our primary hypothesis is that the HRS method automatically and objectively detects and quantifies HIE lesions (in almost real time) comparable to those using manual extraction methods which are currently considered the "gold standard."

METHODS

Bilateral Common Carotid Artery Occlusion With Hypoxia (BCAO-H)

All animal protocols were approved by the Institutional Animal Care and Use Committee. Ischemia was induced by bilateral occlusion of the common carotid arteries in unsexed P10 Sprague-Dawley rat pups (n = 43) while control animals (n = 8) did not experience surgery or hypoxia (8). BCAO-H was induced under anesthesia (4% isoflurane) by a midline neck incision where the common carotid arteries were exposed, separated from the vagal nerves, occluded with 5-0 nylon sutures, and then ligated under a surgical microscope. The wound was sutured and animals were then placed with their dams on a warm heating pad (37°C) for recovery from anesthesia. After 90 minutes BCAO, the pups were further exposed to varying durations of hypoxia (1–15 min; 8% oxygen – 92% nitrogen) to produce graded volumes of injury severity.

MRI

Control and BCAO-H animals underwent MRI at 1 day (examining early injury) and 28 days (approxi-

mating long-term injury) post-injury (8). Pups were lightly anesthetized using isoflurane (1.0%) before imaging and a thermostat-controlled heated water cushion maintained body temperature at 35–37 ± 1°C. Two Bruker Advance MRI scanners (Bruker Biospin, Billerica MA), one at 11.7 Tesla (T) and another at 4.7T were used for acquiring T2-sequence data with 20 coronal slices (1 mm thick and interleaved by 1 mm) (31). At 1 to 13 days, pups were imaged on 11.7T scanner (8.9 cm bore) with a multi-echo T2 sequence with the following parameters (8): TR/TE = 4600 ms/ 10.2 ms, number of echoes = 10, matrix = 128 × 128, field of view (FOV) = 3 × 3 cm with 2 averages, with a total imaging time of 20 min. At 28 days, the maturing pups were imaged on a 4.7T scanner (30-cm bore) with multi-echo T2 sequence with following parameters: TR/TE = 2850 ms/20 ms, number of echoes = 6, matrix = 128 × 128, FOV = 3 × 3 cm with 2 averages, with a total imaging time of 12 min. We observed no variability between the two scanners; furthermore, HRS results use the actual T2 data and their rescaled values (i.e., not arbitrary MRI intensities).

Manual Volumetric Image Analysis

While inter-observer variations in manual lesion quantification result in irreproducibility and subjectivity, intra-observer quantification is typically consistent. Hence, we have used single observer manual detection as the "gold-standard" for comparison. The manual volumetric MR image analysis method in the current study was similar to those previously published (31,32). Using Amira (Mercury Computer Systems, Inc.), T2 images were analyzed, the regions of hyperintensity were used to delineate lesions and the lesion boundaries were manually drawn from which the infarct volume was computed. Data included infarct volumes, noninfarct brain volumes, and total brain volumes. The investigator performing the volumetric measurements was blinded to the hypoxia duration (31). The 2D regions-of-interest (ROIs) on each MRI (i.e., brain slice) were drawn manually and the MR tissue volume of interest was calculated using a 1 mm inter-slice distance.

Hierarchical Region Splitting (HRS)

HRS is an automated and recursive region segmentation method that segments T2 values into uniform image regions recursively. In each recursive splitting, regions from the previous stage (iteration) are separated into smaller yet more uniform image regions (Fig. 1). Different brain tissues have different contrast levels such that a uniform region likely represents a single brain tissue type (e.g., cortex versus striatum). Because T2 images reflect water content and as HII lesions contain more water (edema) than normal brain tissues, the HRS method can exploit these differences to segment uniform T2-regions in which the brightest signal is expected to indicate the location of the brain lesion(s).

The HRS methodological sequence (Fig. 1) to extract lesion and NABM from T2 images are:
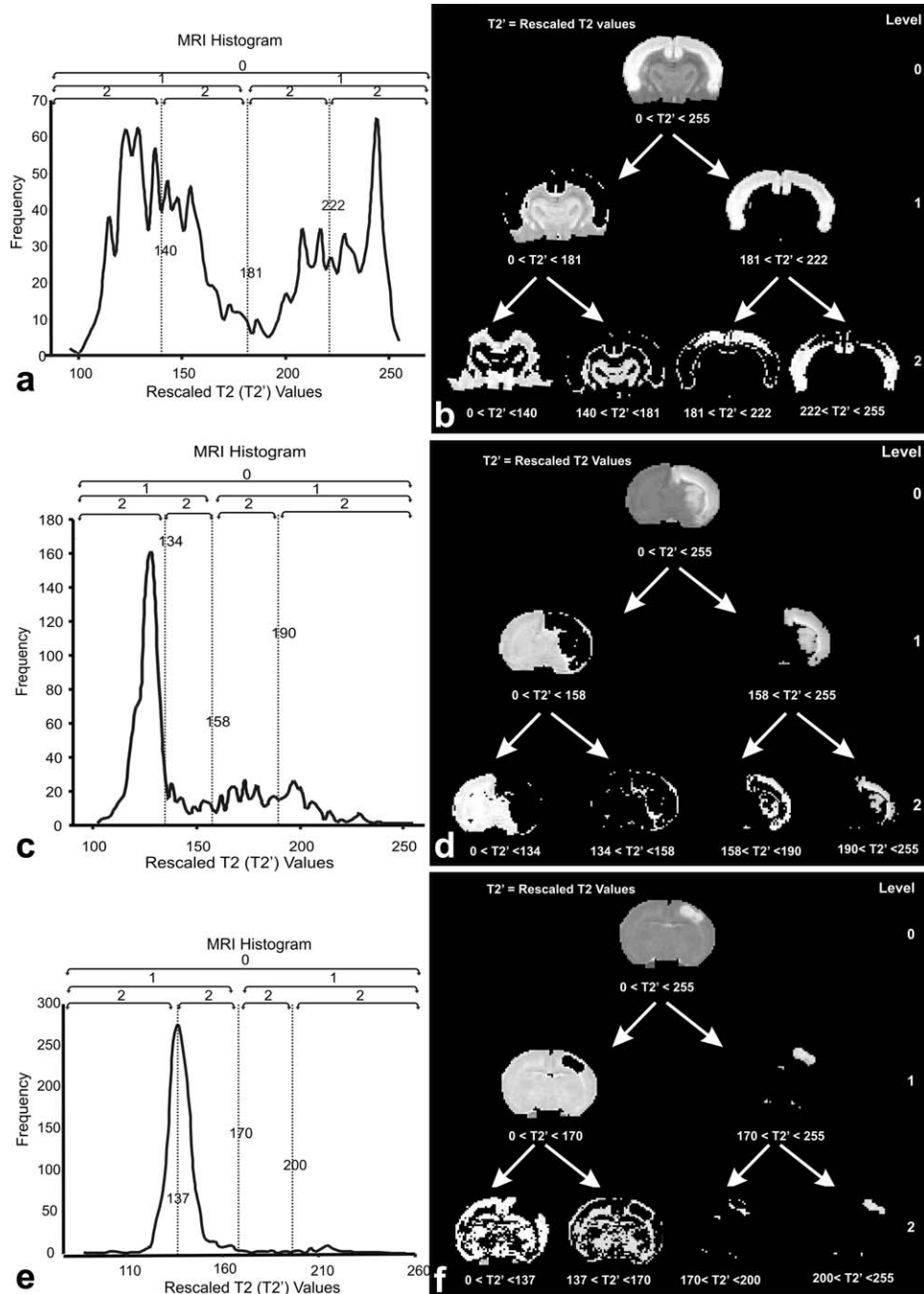
Figure 1. HRS in (a, b) severely, (c, d) moderately and (e, f) mildly injured BCAO-H rat brains. **a,c,e:** The histograms illustrate the global MRI signal spectrum of the rescaled T2 (denoted as T2′ in this study) values in the (a) severe, (c) moderate, and (e) mild BCAO-H animals. **b,d,f:** The automatically detected threshold of (b) T2′ = 181, (d) T2′ = 158, and (f) T2′ = 170 split the histograms and the level 0 images into two parts as shown in level 1. This histogram-based image splitting is repeated and sub-images with intensities from each part of the histogram form the tree structure. Only the top part of the HRS tree is shown where each image represents only a portion of the histogram. HRS is terminated when the split regions are considered uniform based on specified criteria (see methods).

Rescaling

To reduce computational complexity and increase robustness to MR signal variation, we rescaled the T2 values to a range [0, 255] and the scaling factors were saved to map the automatically derived results in image intensity values back to T2 values. As T2 values more than 255 are not common, rescaled data were very similar to the actual T2 values. We have used this rescaling step for compatibility of the HRS method to other MRI modalities (e.g., diffusion weighted images) where the range of values may vary widely.

Deriving Histogram and Distribution Measures

The signal spectrum histogram (H) of the rescaled T2 (T2′ in the range [1,N]) was computed.

$$H(i) \quad i \in [1, N]$$

Histogram H was Normalized to estimate a probabilistic distributed function (p) and cumulative distribution function (Ω), which is cumulative sum of p, was computed.

$$p(i) = H(i) \bigg/ \sum_{i=1}^{N} H(i) \quad \Omega(i) = \sum_{j=1}^{i} p(j)$$

Finally, the cumulative weighted means (μ) and the terminal cumulative weighted mean (μ_t) are computed from probabilistic distribution function p(i).

$$\mu(i) = \sum_{j=1}^{i} p(j)^{*}j \quad \mu_t = \mu(N)$$

Computing Adaptive Segmentation Threshold

Similar to Otsu's method (33), we compute the sequence of Otsu's measure (σ_b^2) using the following equation:

$$\sigma_b^2(i) = \frac{[\mu_t{}^{*}\Omega(i) - \mu(i)]^2}{[\Omega(i)^{*}(1 - \Omega(i))]}$$

Then we find the mode (idx) in the Otsu's measure sequence; if there are more than one (closely spaced) modes, we take the mean of them:

$$\text{idx} = \text{mean}\big[\text{modes}\big(\sigma_b^2\big)\big]$$

The normalized threshold (Th′), which is in between [0,1], is computed by the following equation.

$$\text{Th}' = \frac{(\text{idx} - 1)}{(N - 1)}$$

We modeled the histogram at every level (i.e., level 0–2, Fig. 1b,d,f) as a bimodal distribution with two distinct and distant peaks (Fig. 1a,c,e). We computed the valley between these peaks as an adaptive threshold (Th) to split the image into two sub-region images (for the next level; one region has values greater and the other region has values less than the threshold; Fig. 1b,d,f) that were relatively uniform in intensity. Each peak is a region with a minimum intra-regional and maximum inter-regional T2 variance. For our case, as range is [0,255], that is, N = 255, we find threshold in rescaled value by: Th = Th′ × 255.

Recursive Bimodal Segmentation

A unique aspect of the HRS method is in its recursive application of bi partite segmentation. The sub-images were then recursively further split to generate a tree-like hierarchical data structure (Fig. 1b,d,f). At every recursion, skewness and kurtosis values of the histograms or part of it have been computed by the following equations:

$$\text{Skewness} = \frac{E(x - \mu)^3}{\sigma^3} \quad \text{Kurtosis} = \frac{E(x - \mu)^4}{\sigma^4}$$

where μ is the mean of histogram x, σ is the standard deviation of x, and $E(t)$ represents the expected value of the quantity t.

Criteria for Stopping Segmentation

Recursive splitting was continued until individual segments or sub-regions had near-uniform MRI intensities. Uniformity was measured based on three factors, whether: (a) the individual connected regions were small (area < 50 pixels ≈ 2 mL volume in a single T2 slice) and unlikely to be further partitioned; OR (b) the T2 signal value for the region had a low standard deviation (STD < 10 T2 values), that is, the regional MRI intensities were relatively uniform; AND (c) the T2 histogram for the segmented region had a low kurtosis value (kurtosis < 1.5) where the peak was too distinct to be modeled as a bimodal distribution. The brain regions obtained from HRS (e.g., HRS trees in Fig. 1b,d,f) were not always uniform in T2 (or rescaled T2) values, but as we descend down the HRS tree, we achieve greater uniformity within individual regions. The threshold values or split-stopping criteria (volume < 2 mL; STD < 10; kurtosis < 1.5) were manually selected based on published results (8).

Manually detected lesions had a standard deviation within 20–25 and a mean within 150–255 of the rescaled T2 values (see step 1). A kurtosis threshold of 1.5 detects peaks that are much sharper than Gaussian (normal) distribution (kurtosis = 3) that has been used in other methods including GMM (22). These thresholds do not affect lesion detection but minimize unnecessary oversegmentation of small or very uniform regions which reduces HRS tree size and computational (space and time) complexity.

Detection of HII Lesion

Based on previously published studies (8) as well as our current results using manual detection methods, we determined that a T2 value of 150 can efficiently distinguish the HII lesion from normal tissues. We used this value as a soft threshold, called *threshold for the lesion mean* (meanTh = 150 in rescaled T2 value; see Step 1). We accomplished this by systematically checking the mean value of the sub-regions in the HRS tree one by one, starting from the top (level 0) and gradually descending down branches of the HRS tree until we reached a sub-region with a mean (rescaled) T2 value that is greater than the meanTh (=150). HRS then categorizes this sub-region as an HII lesion (e.g., Fig. 1b, lesion: right-side sub-image in level 1 has a mean rescaled T2 value = 224, while NABM (left-side) sub-image at level 1 has a mean T2 value = 139). Note that, sub-regions were not re-grouped to detect lesion. The threshold − meanTh = 150 was the cut-off for the HII T2 mean. Lesion(s) detected by HRS may have mean (rescaled) T2 values substantially above this meanTh = 150 value, and as a result effective lesion thresholds differ between individual animals.

Finally, regional properties (area, mean, and standard deviation) and signal spectrum or histogram properties (kurtosis and skewness) were computed for the 2D lesions detected by HRS. The 3D volumes were then computed.

Statistical Analysis

We used Student's paired t-test with a two-tailed distribution to determine whether the manual and HRS results were likely to have derived from the same two underlying populations that have the same mean. Lower output probability value (P value) signifies better correlation between results from manual and HRS methods. We also computed correlation coefficients (r^2) of the HRS and manual results to determine how well they can be modeled using linear regression.

Performance Indices

Standard statistical analysis only considers scalar volumetric comparisons without considering overlap between the detected location of the lesions from HRS and manual methods. To compare lesion location(s) and measure performance between the HRS and manual methods, we used three performance indices: (i) sensitivity, (ii) specificity, and (iii) similarity, as defined below (34).

$$Sensitivity = \frac{L_{Man} \cap L_{HRS}}{L_{Man}}$$

$$Specificity = \frac{(BR_{Man} - L_{Man}) \cap (BR_{HRS} - L_{HRS})}{(BR_{Man} - L_{Man})}$$

$$Similarity = 2 \times \frac{L_{Man} \cap L_{HRS}}{L_{Man} \cup L_{HRS}}$$

where L_{Man} and L_{HRS} were the binary lesion masks derived from manual detection and HRS, respectively, and where BR_{Man} and BR_{HRS} were the binary masks for the total brain as derived from the manual and HRS detection methods, respectively. Sensitivity defined how well the lesions overlap between these two methods and quantified true positive detection. Specificity quantified overlap in NABMs from both methods and signified the negative detection rate. The similarity index measured location and volume similarities between lesions derived from both methods. Note that the similarity index actually measures the accuracy of HRS detection in comparison to the gold-standard, in terms of location, size and overlap of the detected HII lesion. Sensitivity and specificity may range between [0, 1], while similarity can vary between [0, 2].

RESULTS
Animals and Neuroimaging

All animals that survived 3 days (81.8%) post-BCAO-H survived until study completion. Mortality rates were higher in the more severely injured pups: mild (i.e., lesion < 15% total brain) was 5.3%, moderate (15% <
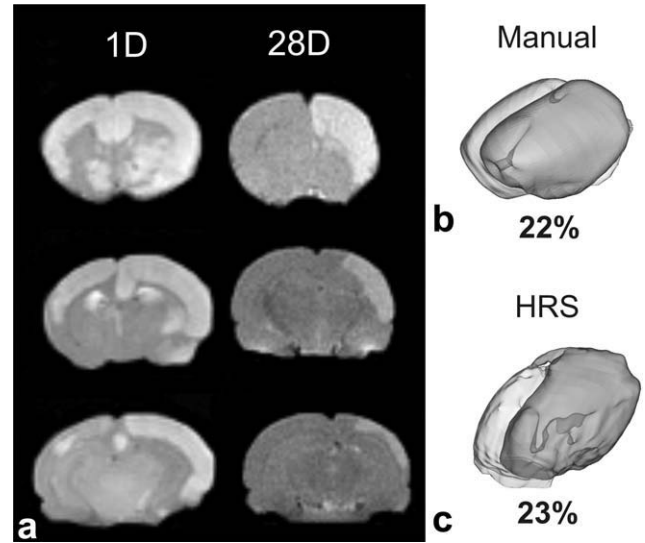

Figure 2. MR volumetric analysis of BCAO-H lesions. **a:** T2 weighted images in a moderately injured BCAO-H rat brain at anterior (top row), middle (middle row), and posterior (bottom row) brain levels. Images at 1 day (left column) and 28 days (right column) illustrate early and long-term injury. **b:** Manually derived 3D T2 reconstructed brain (light-gray) and lesion (dark-gray) volumes at 1 day. **c:** HRS-derived 3D brain (light-gray) and lesion (dark-gray) volumes at 1 day. Note similarity between manual (22%) and HRS (23%) - derived volumes.

lesion < 35%) was 20%, and severe (lesion > 35%) was 44%.

Lesion distribution (1 day post BCAO-H) was dependent upon severity (Fig. 2a). Animals with mild injury primarily exhibited unilateral cortical infarctions, those with moderate injury had bilateral cortical involvement, whereas those with severe injury had more extensive bilateral and in some cases hippocampal involvement. At 28 days, similar injury patterns were observed but were often smaller (Fig. 2a) (8).

Manually Derived Lesion Volumes

Manually extracted 3D volumes at 1 day were 0.7 ± 2.3%, 26.3 ± 6.2%, and 48.7 ± 5.5% for mild, moderate and severe lesions, respectively (Table 1). At 28 days, all severe animals recovered to become moderate/mild and hence there were no severe animals (i.e., lesion > 35%). An example of the 3D volumetric representation of a manually extracted moderate lesion and NABM at 1 day is depicted in Figure 2b.

HRS-Derived Lesion Volumes

HRS has lesions in mild (Fig. 1e,f), moderately (Fig. 1c,d) and severely (Fig. 1a,b) injured BCAO-H animals, from which we derived the 3D lesion volumes (Fig. 2c). At 1 day after BCAO-H, mean ± standard deviation of HRS-detected percentage lesion volumes (with respect to the entire brain) were for mildly injured: 5.95 ± 8.4%, for moderately: 30.4 ±

Table 1
Percent Volume of Ischemic Injury: Comparison of Manual Versus HRS Methods

	Manual	HRS	r^2 value	*P* value
MRI 1 d after HII				
Mild	0.73 ± 2.3	5.95 ± 8.35	0.49	0.0166
Moderate	26.25 ± 6.2	30.40 ± 5.50	0.99	0.0086
Severe	48.74 ± 5.5	52.98 ± 10.10	0.80	0.0571
MRI 28 d after HII				
Mild	0.00 ± 0.00	5.54 ± 4.70	—	0.0004
Moderate*	21.18 ± 6.52	26.53 ± 8.73	1.00	0.1803
Severe**	—	—	—	—

*These values exclude a single outlier data point from the 28 d group that led to significant difference between manually (22.8%) and HRS (45.8%) derived percentage lesion. Inclusion of this outlier results in manually detected percentage injury volume to be (21.72 ± 4.70) and that from HRS (32.94 ± 12.70) with a correlation of $r^2 = 0.65$; *P* = 0.1995.
**All severe animals had recovered by 28d and became either moderate or mild. Hence no severe animal results could be reported for 28d.

5.5%, and for severely injured animals: 52.98 ± 10.1%. Mean and standard deviation for the same injury subgroup remained nearly the same for 1 day and 28 days when comparing mild and moderate animals (Table 1).

Comparing HRS and Manual Lesion Volumes

HRS-derived lesions in mild, moderate, and severely lesioned animals (Fig. 3) were consistently larger than those derived manually (Fig. 3). There was a 1–4% increase in lesion volume with HRS (Table 1). We found a very strong correlation (Fig. 4) for all animals between lesion volumes as measured by both methods ($r^2 = 0.95$; $P = 8.6 \times 10^{-7}$). Likewise, when we examined the injury severity subgroups, we found a strong correlation between manual and HRS-derived percentage volumes as follows: mild ($r^2 = 0.49$; $P = 0.0166$), moderate ($r^2 = 0.99$; $P = 0.0086$) and severe ($r^2 = 0.80$; $P = 0.0571$) (Table 1). For NABM, 3D volumes computed by manual and HRS methods were also strongly correlated ($r^2 = 0.92$; $P = 1.4 \times 10^{-22}$).

Performance indices (sensitivity, specificity, and similarity) for locational accuracy of HRS based detection compared with manual detection method were consistently high (sensitivity: 0.82 ± 0.15; specificity: 0.86 ± 0.12; similarity: 1.47± 0.40) for all 43 BCAO-H animals that we assessed (Table 2; Fig. 5b).

DISCUSSION

Although described in adult stroke (11), automated HII lesion detection has not been reported in human newborns nor in animal models of neonatal HII. The neonatal brain has different physical characteristics than the adult and the response to injuries such as HII are also different. The neonatal brain has more water content, smaller brain volume, less myelination and anatomical definition (8), and most importantly, more heterogeneity within brain and lesion data due to ongoing nervous system maturation. Due to these complexities, automated lesion detection methods utilized in adults, like those using established anatomical brain maps (24), GMM (17) or other standard probabilistic models (28), are not feasible for neonatal

HII detection. The principal findings of the current study are that: (i) HRS is a mathematically less intensive, yet a powerful automated approach to detect, localize and quantify T2 MRI lesions in a neonatal model of HII; and (ii) HRS measurement of neonatal HII lesion volumes are similar to the currently used "gold standard" of manual quantification with its advantages being speed, accuracy and robustness. The speed of HRS (15 s per brain) is several orders of magnitude faster than the manual method (2–3 h per brain) and as such for both experimental work and
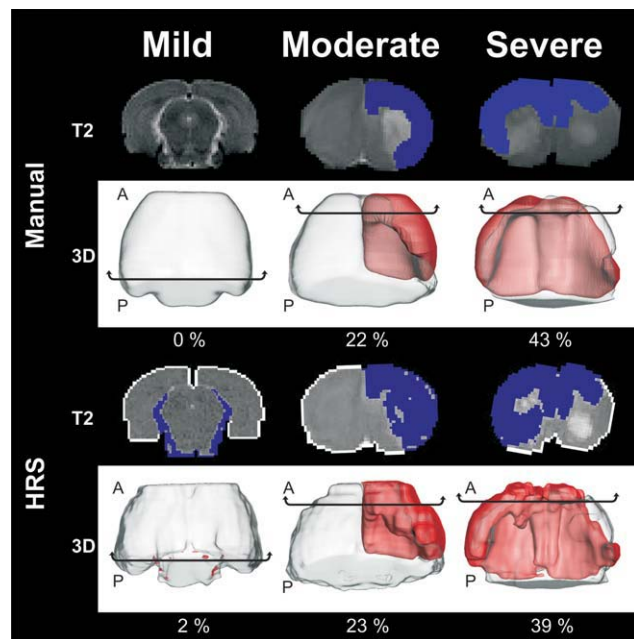

Figure 3. Comparative volumetric results between HRS and manual methods at different injury-severities. Manually detected lesions in 2D and in 3D from T2WI and automatically (HRS) detected lesions in 2D and in 3D from T2WI are similar. T2WI from mild (<15% lesion), moderate (15–35%), and severe (>35%) injuries are shown at the level of the horizontal line. The percentage of the lesion volume compared with the entire brain is shown below the 3D volumes. Comparing the percentage-lesion volumes, and 3D volumetric results, we demonstrate that HRS values are remarkably similar to those determined for the manually detected lesions.
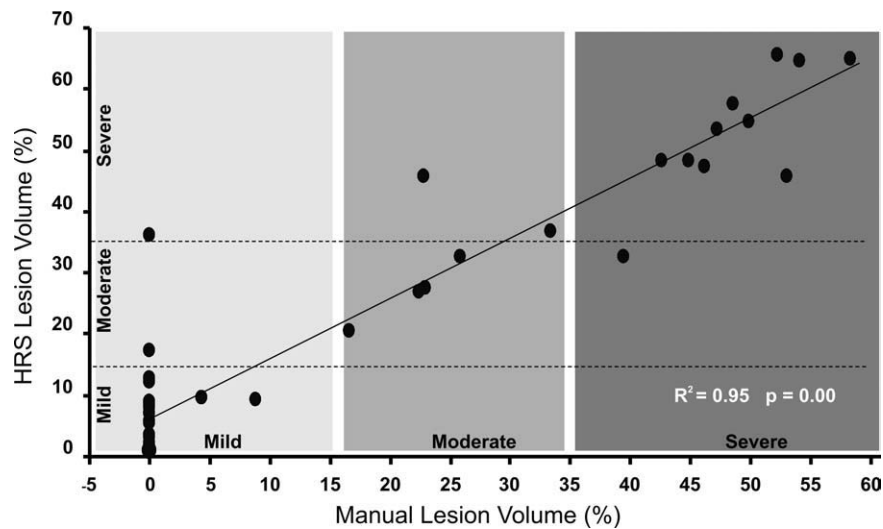
Figure 4. Statistical comparisons between manual and HRS lesion volumes. Cross-correlation of manual percentage lesion volumes compared with the HRS-detected lesion volumes. The correlation factor $r^2 = 0.95$ at ($P = 8.6 \times 10^{-7}$) demonstrates that HRS performs similarly to manual detection methods. Most importantly, in most cases, the BCAO-H injured brains were classified at the same severity levels based on percentage lesion volumes, mild (lesion < 15%), moderate (15% < lesion < 35%), and severe (lesion > 35%) in both types of analysis.

eventual clinical use offers significant practical advantages. HRS is more accurate than the manual method in the sense that it can detect those lesions which have T2 values high enough to be considered as lesion but which the human eye may overlook due to smaller size or relative low contrast with respect to the neighboring voxels. Human visual perception suffers from inconsistencies in processing image data and because HRS is an objective consistent method, it can accurately detect these lesions. The robustness of HRS is evident from the consistency (low standard deviations) of our results (Fig. 4) and the performance indices (Fig. 5b) compared with the manual methods from 43 rat brains with different injury severities, different affected anatomical locations (Figs. 1, 3), and different imaging time-points. As seen from the histogram plots in Figure 1a (severe injury), Figure 1c (moderate injury) and Figure 1e (mild injury), the peaks for NABM and lesion are not always clearly demarcated. With larger lesion size and improved MR contrast, the lesion peaks appear more prominent. However, irrespective of lesion percentage volume, HRS detects and quantifies the lesion objectively. Our findings also suggest that the HRS framework could be used clinically for lesion detection and volume. We acknowledge that before translating the HRS method

to clinical applications, we have to consider how to adapt HRS to account for the anatomical location of the lesion(s).

Although Otsu's method (33) has been used for adaptive image segmentation, to the authors' knowledge it has not been applied recursively for hierarchical region splitting of MRI. A recent publication (35) demonstrated a histogram and clustering based adult TBI detection method that worked well only for MR slices with visually detectable lesions (i.e., biased). Our HRS method is more generic and unbiased, and broadens the horizon of automated lesion detection. Although structurally HRS-derived trees appear to look like the decision trees (DT) used in machine learning systems, they are fundamentally different. A single DT is developed from supervised learning of a training data set and each node acts as rule-base for an expert system. In contrast, with HRS trees, there is no learning/ training phase and it produces completely different hierarchical trees for each MRI slice in an unsupervised, adaptive and data-driven way. The nodes in HRS trees do not classify region splitting results, as this is done by an external parameter (meanTh in lesion detection). The novelty of our approach is in the use of recursive splitting (without any merging) to get a multi-level representation of the MRI data (i.e., different abstraction or granularity

Table 2

Sensitivity, Specificity, and Similarity Scores by Injury Severity Levels

	Mild	Moderate	Severe	Overall
Sensitivity	0.47 ± 0.12	0.86 ± 0.08	0.85 ± 0.09	0.82 ± 0.15
Specificity	0.87 ± 0.02	0.86 ± 0.18[a]	0.87 ± 0.07	0.86 ± 0.12
Similarity	0.40 ± 0.25	1.53 ± 0.13	1.63 ± 0.17	1.47 ± 0.40

[a]Note that, for skewed histograms (or distributions), as happened to be in the case of moderately injured brains (specifically due to one single brain, where specificity is as low as 0.42 compared to the others above 0.90), we obtained a high standard deviation and (mean + SD) become higher than the upper-limit of the values (as happened for specificity of the moderate brains).
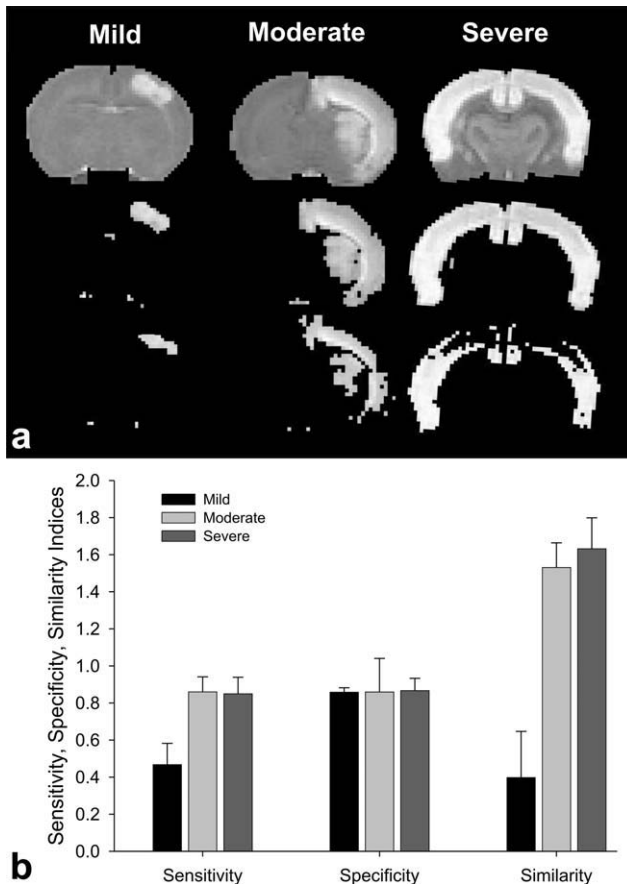
Figure 5. a: Comparison of HRS performance for BCAO-H at different injury severity levels. Each column is the right-most branch in HRS trees (in Fig. 1b,d,e) for corresponding animals. HRS could detect different regions and locations of the lesions (middle row) from the entire brain (top row). Even inside the lesion, most critically affected tissues (with much higher T2 values) could be detected (bottom row). **b:** Performance measures to quantify locational accuracy of HRS lesion detection compared with manual lesion detection. Sensitivity (range between [0, 1]) measures the true positive detection rate while specificity (range between [0, 1]) measures the true negative detection rate. Similarity (range between [0, 2]) measures size and location-specific reliability. The graph illustrates that, for moderate and severely injured animals, HRS detects lesions of a similar size and at the same locations compared with manual methods. In the case of mildly injured animals, HRS is not as robust in detecting small lesions at different locations (sensitivity and similarity) compared with the manual approach (see the Discussion section).

levels) and application to medical image processing, where these methods have not been previously used. HRS oversegmentation does not affect detection and its usefulness is in detecting heterogeneity *inside* any sub-regions within the lesion to characterize different patterns of injury (see Fig. 5a) for better outcome prediction (future work), as the MR values are not uniform in detected lesions.

Computationally intensive mathematical models are generally used for brain segmentation but their complexity hampers rapid detection, localization (i.e., detected location in brain MRI) and volumetric quantification of lesions (12,14,22,27). HRS provides a robust solution in that it is simple, fast, and yet is a reliable automated lesion detection method. HRS linearly maps any MR value (independent of modality) to a manageable image intensity range, detects the lesions and allows mapping of the results back to the original relevant MRI values. HRS worked well (high correlation with manually detected results) for 43 brains with different noise-levels (from different MRI machines), different lesion volumes (mild, moderate, and severe injury levels) and over two broad imaging time-points (1 day and 28 days after HIE). We believe that HRS, with our rescaling strategy, worked well. It uses the rescaled image for regional intensity properties to detect internal details at an even finer level within the MRI. Successive splitting of HRS images are performed using automatically and adaptively selected thresholds from the histogram-based MR signal spectrum at different hierarchical levels. Because different brain regions contrast differently, the HII lesion and NABM can also be automatically and adaptively detected.

HRS is a top-down locally adaptive automated image segmentation method; top-down as we start from the entire MRI data spectrum (histogram), and local because it considers the histogram characteristic of each individual region in the hierarchy of segmentation levels. HRS is adaptive to the characteristic of local regions. The derived threshold values are not affected by small MRI signal variations because HRS maximizes the difference in the means of the two regions while minimizing the variance within each region. It is capable of detecting lesions despite different degrees of injury severity or location of the lesion in different brain anatomical regions (e.g., cortex, hippocampus, etc.) (Table 1; Figs. 3, 5a). Most importantly, it correctly classifies the degree of injury severity based on the percentage lesion volume in most cases (Fig. 4) when compared with the manual "gold-standard" approach.

The majority of studies on MRI based lesion detection in humans and animals have not considered automated segmentation (31,32,36–39). Manual lesion detection by a single trained individual is remarkably good as evidenced by intra-observer studies. However, in order for lesion detection to be clinically relevant, a broad range of individuals (with different backgrounds and training levels) would have to be deployed to make this viable. Thus, inter-observer variability is and would be significant. Besides their subjective nature, manual methods are also time consuming (2–3 h/brain), which can lead to fatigue-related errors. Thus, manual results between individuals are not comparable for statistical comparisons. In contrast, HRS is fast (10–15 s/brain), objective and robust. Using manual quantification as the "gold standard", HRS detects the same anatomical location of the lesion in all animals examined (Fig. 5b) with very strong volumetric correlations (Fig. 4). For animals with mild brain injury, HRS overestimated the percentage of lesion volume compared with manually derived results (0% in most cases). This likely occurred because, in manual methods, the human visual system "overlooks" smaller lesion regions (as expected with mild BCAO-H) due to the visual

illusions caused by neighboring pixel data (9). Also lesion volumes and their 3D reconstructions from both methods (Figs. 2b,c, Fig. 3) demonstrate that HRS-derived volumes are less smooth than those from the manual methods due in part to human bias toward smoothing surfaces. Automated methods have their own limitations and pitfalls but most studies (34) have compared their automatically derived results to manually derived ones. Specifically, as manual results from a single trained observer have low intra-observer variability such manual results can be used as a "gold standard". In the study by Shen and colleagues (34) sensitivity (mean, 0.96), specificity (mean, 0.99) and similarity (mean 0.97) were computed only for 15 MR images and visually detectable lesions and similarity were quite low for its range [0,2]. This study also ignored other MR slices where false-positives were expected; hence, the specificity result reported is biased. Finally, this study used a simulated lesion that was spherical, though lesions typically are not spherical. In contrast, our study used a large and heterogeneous dataset (43×12 MR images without bias) and reports performance scores (Table 2) comparable to, if not better than, published results.

Note that histograms of MRI slices at different injury severity levels may look significantly different (mild: Fig. 1e; moderate: Fig. 1c; severe: Fig. 1a), but HRS can automatically define the best threshold (T2 = 170 (mild), 158 (moderate), 181 (severe) in our examples) for segmentation of the corresponding histograms. Finally, HII lesion detection by HRS is also adaptive, in terms of the following: (i) level in the HRS tree, (ii) subimage in the HRS tree, and (iii) T2 values (and their mean) of the pixels in the detected region (with the only constraint being that mean T2 < meanTh = 150).

It is acknowledged that without histological confirmation, it cannot be determined equivocally whether the manual detection or the HRS method is absolutely correct for the dynamically evolving lesion. However, our previously published work (40) demonstrated that the manual lesion determination from MRI correlated well with histology at later time points (28 days), but underestimated the lesion volume at early time points (3 days). This fact supports that HRS-derived lesion volumes appear to be larger than manually detected lesions at 1 day. This confirms that HRS and manual methods are comparable, but that HRS detection is much faster and automated.

In its current form, two limitations of the HRS approach are recognized (and correctable). (i) HRS considers the global spectrum (histogram) of individual MRI slices and can occasionally wrongly detect small lesions (false-positives). Future work will use 2D/3D continuity and proximity to reject such false-positives. (ii) HRS also falsely detected lesions at the level of the brainstem (Fig. 3) that led to ~4% larger volumes than those from the manual method. This is likely due to bright T2 CSF signal, but use of Fluid Attenuated Inversion Ratio (FLAIR) images would minimize these errors. Also, after lesion detection (unlike in (24)), anatomical atlases and published MR (e.g., T2) values for different brain tissues can be used in a model-based rejection of outliers.

In conclusion, HRS is fast, robust and capable of quantifying HII lesion volume. HRS fulfills the current need to rapidly and objectively quantify injury volume(s) from neuroimaging. HRS has a broad applicability to other types of acquired brain injury (e.g., trauma, stroke). Only the global threshold (meanTh) would be different for different applications and different types of injuries or other tracked tissues (e.g., stem cells). The ability to rapidly segment and quantify volumes of injury is instrumental in using MRI in translational research for candidate selection for treatment. Lesion shape considerations and 3D connectivity based outlier rejection will further improve our results and we will consider these in our future studies. HRS also can be used serially to monitor the evolution of disease over time and to objectively quantify the benefits (or lack thereof) of specific treatments. Thus, the HRS approach has the potential to define early biomarkers of treatment successes/failures that could trigger clinical choices to adjust therapies to improve success.

REFERENCES

1. Taeusch HW, Ballard RA, Gleason CA, Editors. Avery's diseases of the newborn. 8th ed. Philadelphia: Elsevier Saunders; 2005.
2. Yager JY, Ashwal S. Animal models of perinatal hypoxic-ischemic brain damage. J Pediatr Neurol 2009;40:156–167.
3. Barnett A, Mercuri E, Rutherford M, et al. Neurological and perceptual-motor outcome at 5 - 6 years of age in children with neonatal encephalopathy: relationship with neonatal brain MRI. Neuropediatrics 2002;33:242–248.
4. Fraser M, Bennet L, Helliwell R, et al. Regional specificity of magnetic resonance imaging and histopathology following cerebral ischemia in preterm fetal sheep. Reprod Sci 2007;14:182–191.
5. Yang J, Wu EX. Detection of cortical gray matter lesion in the late phase of mild hypoxic-ischemic injury by manganese-enhanced MRI. Neuroimage 2008;39:669–679.
6. Barkovich AJ, Hajnal BL, Vigneron D, et al. Prediction of neuromotor outcome in perinatal asphyxia: evaluation of MR scoring systems. AJNR Am J Neuroradiol 1998;19:143–149.
7. Haataja L, Mercuri E, Guzzetta A, et al. Neurologic examination in infants with hypoxic-ischemic encephalopathy at age 9 to 14 months: use of optimality scores and correlation with magnetic resonance imaging findings. J Pediatr 2001;138:332–337.
8. Recker R, Adami A, Tone B, Tian HR, Lalas S, Hartman RE, Obenaus A, Ashwal S. Rodent neonatal bilateral carotid artery occlusion with hypoxia mimics human hypoxic-ischemic injury. J Cereb Blood Flow Metab 2009;29:1305–1316.
9. Niimi T, Imai K, Maeda H, Ikeda M. Information loss in visual assessments of medical images. Eur J Radiol 2007;61:362–366.
10. Shapiro L, Stockman G. Computer vision, Chapter 10: Image segmentation. Upper Saddle River, NJ: Prentice Hall, Inc; 2001. p 279–325.
11. Kabir Y, Dojat M, Scherrer B, Forbes F, Garbay C. Multimodal MRI segmentation of ischemic stroke lesions. In: Proceedings of the 29th Annual International Conference on the IEEE Engineering in Medicine and Biology Society, (EMBS), Lyon, France. 2007. p 1595–1598.

12. Bricq S, Collet C, Armspach JP. Unifying framework for multimodal brain MRI segmentation based on hidden Markov chains. Med Image Anal 2008;12:639–652.

13. Castellano G, Bonilha L, Li LM, Cendes F. Texture analysis of medical images. Clin Radiol 2004;59:1061–1069.

14. Liao L, Lin TS, Li B. MRI brain image segmentation and bias field correction based on fast spatially constrained kernel clustering approach. Pattern Recognit Lett 2008;29:1580–1588.

15. Nakamura K, Fisher E. Segmentation of brain magnetic resonance images for measurement of gray matter atrophy in multiple sclerosis patients. Neuroimage 2009;44:769–776.

16. Fan X, Yang J, Zheng YJ, Cheng LS, Zhu Y. A novel unsupervised segmentation method for MR brain images based on fuzzy methods. Lect Notes Comput Sci 2005;3765:160–169.

17. Greenspan H, Ruf A, Goldberger J. Constrained Gaussian mixture model framework for automatic segmentation of MR brain images. IEEE Trans Med Imaging 2006;25:1233–1245.

18. He Q, Karsch K, Duan Y. A novel algorithm for automatic brain structure segmentation from MRI. Lect Notes Comput Sci 2008; 5358/2008:552–561.

19. Song Z, Tustison N, Avants B, Gee J. Adaptive graph cuts with tissue priors for brain MRI segmentation. In: Proceedings of the 3rd IEEE International Symposium on Biomedical Imaging (ISBI), Arlington, Virginia, 2006. p 762–765.

20. Ibrahim M, John N, Kabuka M, Younis A. Hidden Markov models-based 3D MRI brain segmentation. Image Vis Comput 2006; 24:1065–1079.

21. Rouainia M, Medjram MS, Doghmane N. Brain MRI segmentation and lesions detection by EM algorithm. World Academy of Science, Engineering and Technology, 2006;24:139–142.

22. Freifeld O, Greenspan H, Goldberger J. Lesion detection in noisy MR brain images using constrained GMM and active contours. In: Proceedings of the IEEE International Symposium on Biomedical Imaging, (ISBI), Arlington, Virginia, 2007;596–599.

23. Bricq S, Collet C, Armspach JP. Markovian segmentation of 3D brain MRI to detect multiple sclerosis lesions. In: Proceedings of the 15th IEEE International Conference on Image Processing (ICIP), San Diego, California, 2008;733–736.

24. de Boer R, van der Lijn F, Vrooman HA, Vernooij MW. Automatic segmentation of brain tissue and white matter lesions in MRI. In: Proceedings of the International Symposium on Biomedical Imaging, (ISBI), Arlington, Virginia, 2007;652–655.

25. Dugas-Phocion G, Gonzalez MA, Lebrun C, Chanalet S, Bensa C, Malandain G, Ayache N. Hierarchical segmentation of multiple sclerosis lesions in multi-sequence MRI. In: Proceedings of IEEE International Symposium on Biomedical Imaging, (ISBI), Arlington, Virginia, 2004;1:157–160.

26. Antel SB, Collins DL, Bernasconi N, et al. Automated detection of focal cortical dysplasia lesions using computational models of their MRI characteristics and texture analysis. Neuroimage 2003; 19:1748–1759.

27. Loyek C, Woermann FG, Nattkemper TW. Detection of focal cortical dysplasia lesions in MRI using textural features. Bildverbeitung fur die Medizine 2008, Informatik Aktuell. Berlin Heidelberg: Springer; 2008, p 432–436.

28. Agam G, Weiss D, Soman M. Probabilistic Brain Lesion Segmentation In DT-MRI. In: Proceedings of IEEE International Conference on Image Processing (ICIP), Atlanta, Georgia, USA; 2006;89–92.

29. Dokladal P, Bloch I, Couprie M, Ruijters D, Urtasun R, Garnero L. Topologically controlled segmentation of 3D magnetic resonance images of the head by using morphological operators. Pattern Recognit 2003;36:2463–2478.

30. Prastawa M, Gerig G. Brain lesion segmentation through physical model estimation. Book Series. Lect Notes Comput Sci 2008; 5358/2008:562–571.

31. Ashwal S, Tone B, Tian HR, Chong S, Obenaus A. Comparison of two neonatal ischemic injury models using magnetic resonance imaging. Pediatr Res 2007;61:9–14.

32. Ashwal S, Holshouser BA, Tong KA. Use of advanced neuroimaging techniques in the evaluation of pediatric traumatic brain injury. Dev Neurosci 2006;28:309–326.

33. Otsu N. A Threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 1979;9:62–66.

34. Shen S, Szameitat AJ, Sterr A. Detection of infarct lesions from single MRI modality using inconsistency between voxel intensity and spatial location–a 3-D automatic approach. IEEE Trans Inf Technol Biomed 2008;12:532–540.

35. Hillary FG, Biswal BB. Automated detection and quantification of brain lesions in acute traumatic brain injury using MRI. Brain Imaging Behav 2009;3:111–122.

36. Barkovich AJ, Miller SP, Bartha A, et al. MR imaging, MR spectroscopy, and diffusion tensor imaging of sequential studies in neonates with encephalopathy. AJNR Am J Neuroradiol 2006;27: 533–547.

37. Saunders DE, Clifton AG, Brown MM. Measurement of infarct size using MRI predicts prognosis in middle cerebral artery infarction. Stroke 1995;26:2272–2276.

38. Schiemanck SK, Post MW, Kwakkel G, Witkamp TD, Kappelle LJ, Prevo AJ. Ischemic lesion volume correlates with long-term functional outcome and quality of life of middle cerebral artery stroke survivors. Restor Neurol Neurosci 2005;23:257–263.

39. Vannucci RC, Vannucci SJ. Perinatal hypoxic-ischemic brain damage: evolution of an animal model. Dev Neurosci 2005;27: 81–86.

40. Ashwal S, Tone B, Tian HR, Chong S, Obenaus A. Serial magnetic resonance imaging in a rat pup filament stroke model. Exp Neurol 2006;202:294–301.