Automated Bayesian Segmentation of Microvascular White-Matter Lesions in the ACCORD-MIND Study

Herskovits EH1*, Bryan RN1, Yang F2

1 Department of Radiology, University of Pennsylvania, Philadelphia, USA 2 Department of Biomedical Engineering, University of California, Irvine, USA

* CORRESPONDING AUTHOR: Department of Radiology, University of Pennsylvania, 3600 Market Street, Suite 370, Room 117, Philadelphia, PA 19104, USA telephone: 215 615 3705; fax: 215 615 3681 e-mail: ehh@ieee.org (Edward H Herskovits)

Received 11.06.2008 Accepted 05.08.2008 Advances in Medical Sciences Vol. 53(2) · 2008 · pp 182-190 DOI: 10.2478/v10039-008-0039-3 © Medical University of Bialystok, Poland

ABSTRACT

Purpose: Automatic brain-lesion segmentation has the potential to greatly expand the analysis of the relationships between brain function and lesion locations in large-scale epidemiologic studies, such as the ACCORD-MIND study. In this manuscript we describe the design and evaluation of a Bayesian lesion-segmentation method, with the expectation that our approach would segment white-matter brain lesions in MR images without user intervention.

Materials and Methods: Each ACCORD-MIND subject has T1-weighted, T2-weighted, spin-density–weighted, and FLAIR sequences. The training portion of our algorithm first registers training images to a standard coordinate space; then, it collects statistics that capture signal-intensity information, and residual spatial variability of normal structures and lesions. The classification portion of our algorithm then uses these statistics to segment lesions in images from new subjects, without the need for user intervention. We evaluated this algorithm using 42 subjects with primarily white-matter lesions from the ACCORD-MIND project.

Results: Our experiments demonstrated high classification accuracy, using an expert neuroradiologist as a standard.

Conclusions: A Bayesian lesion-segmentation algorithm that collects multi-channel signal-intensity and spatial information from MR images of the brain shows potential for accurately segmenting brain lesions in images obtained from subjects not used in training.

Key words: Computer-Assisted Radiographic Image Interpretation, Bayesian Analysis, Brain, Magnetic Resonance Imaging

INTRODUCTION

White-matter lesions (WMLs) are common findings in magnetic-resonance (MR) brain examinations of elderly subjects [1-3]; they are usually caused by diseases, such as hypertension and diabetes [4]. MR techniques manifest WMLs because of their increased tissue-water content, or loss of myelin. On T1-weighted images, lesions appear hypointense or isointense relative to normal brain, and on T2-weighted, spin-density (SD)-weighted, and FLAIR images, these lesions appear hyperintense. Because of this high tissue contrast, large-scale epidemiological studies of cardiovascular risk factors have increasingly relied on MR examination to determine the effects of these conditions on the brain.

Manual segmentation is the most common way to delineate WMLs, but this process is very time consuming, and

may suffer from poor inter-rater reliability. To simplify this process, some researchers involved in studies of large cohorts have adopted simplified assessment scales for brain lesions [2,5], which provide only regions (e.g., "infratentorial"), rather than coordinates, to describe lesion locations. Furthermore, segmentation results may differ among raters, or even for the same rater at different times; automated techniques promise greater reliability [6,7]. Such considerations are particularly relevant to longitudinal studies involving hundreds, or perhaps thousands of subjects, such as the Cardiovascular Health Study [2,5] or the ACCORD-MIND study [8].

Despite the potential advantages of automated brain-lesion segmentation, manual segmentation continues to be favored, due to the difficulty of automatically segmenting brain lesions, as well as the lack of software that can be adapted to a variety of lesions. To address these challenges of automation and generalizability, researchers have proposed many different brain-lesion segmentation algorithms [6,7,9-15]. These methods can be grouped into semi-automatic and fully automatic classification methods. Algorithms that require nontrivial user intervention (i.e., mouse operations, or the specification of more than a few parameters) [6,7,12] are semi-automatic. Although many of these methods are less time-consuming and are more reliable than manual segmentation, even semi-automatic methods may not be scalable to studies involving thousands of subjects.

For the segmentation of large numbers of images, fully automatic methods [11,13-15] are required; however, some of these approaches may still require nontrivial user input. For example, Anbeek et al. used a K-nearest neighbor classifier to segment lesions [15], and van Leemput et al. segmented MS lesions as outliers of normal-tissue intensity models [13]. They performed intensity-based brain-tissue classification, in which those voxels that are not well classified by these intensity models are identified as lesions. In van Leemput's method, there is a user-determined parameter κ , which affects the segmented lesion yield: larger k yields more segmented lesions, at the risk of a higher false-positive rate. Admiraal-Behloul described a method based on clustering and reasoning [14]. This approach includes ad-hoc techniques to reduce the false-positive rate; they allow the user to delineate manually regions in the template in which lesions cannot occur.

Our aim is to design, implement and evaluate a fully automatic brain-lesion-segmentation method suitable for the processing of image data from thousands of subjects. Toward this end, we propose a novel segmentation algorithm based on Bayesian methods for combining multivariate signal-intensity and spatial information. Our method is supervised; we use data from a training set to determine classification statistics. In contrast to most other segmentation methods [11,13-15], our approach is fully automatic; it uses only these statistics to classify new subjects. We performed experiments to determine whether our method detects brain lesions accurately, runs efficiently, and is robust to scanner variability, which is vital for multi-site trials. We tested our approach using data from the ACCORD-MIND study, a prospective study of adult diabetics with minimal or no baseline neurological disease; thus, we expected that most subjects would have few or no brain lesions, and that any brain lesions found would be vascular (e.g., microvascular white-matter lesions, or old lacunar strokes).

MATERIALS AND METHODS

Materials

This research was approved by the University of Pennsylvania Institutional Review Board, and by the ACCORD-MIND review board. We obtained MR images of the head, including T1-weighted, T2-weighted, spin-density-weighted and FLAIR sequences, for 42 subjects at either of two sites from the ACCORD-MIND study: 30 examinations had been acquired from Hennepin County Medical Center, Minneapolis, MN, using a 1.5T Philips Intera instrument (site 1), and 12 examinations had been acquired at the Winston-Salem, NC site, using a 1.5T General Electric LX instrument (site 2). The resolution of the T1 sequence was 0.9375 mm \times 0.9375 mm \times 1.5 mm; for other sequences, the resolution was 0.9375 mm \times 0.9375 mm \times 0.9375 mm \times 3 mm. Field of view was 240 mm for all sequences. The acquisition parameters for the Philips instrument were (TR/TE/TI, ms): T1: 21/8, T2: 2630/100, SD: 2630/27, FLAIR: 8000/100/2400; those for the General Electric instrument were: T1: 24/8, T2: 3200/122, SD: 3200/30, FLAIR: 8002/100/2000.

To provide a ground-truth data set, a senior neuroradiologist (RNB) manually labeled brain lesions for all subjects, relying primarily on FLAIR images, but with access to all sequences for each subject. Criteria for lesion identification were similar to those used in the Cardiovascular Health Study [2, 5].

Segmentation Method Theory and Assumptions

For multi-channel image data, we use (L_v, I_v) to denote the location and the *n*-channel signal-intensity information of a voxel v, respectively. In order to capture normal anatomic variation, we applied an elastic-registration algorithm [16,17] to transform each subject's MR images to a standard coordinate space *D*; we refer to this step as *spatial normalization*.

Given the registered multi-channel image data O, our task is to classify each voxel into one of the tissue classes, yielding the segmented image S. According to Bayes' theorem, we obtain the optimal segmentation S^* by maximizing P(S | O): $S^* = \arg \max P(S | O)$.

To determine which statistics to collect from the training set, we make the following assumptions: **Assumption 1** The classification *C* of each voxel is independent of the classification of every other voxel. Thus,

$$\arg\max_{S} P(S \mid O) = \bigcup_{v} \arg\max_{C} P(v \in C \mid I_{v}, L_{v}).$$
(1)

This assumption allows us to optimize the classification of each voxel independently, greatly simplifying the algorithm, at the expense of ignoring interactions among voxels. By Bayes' rule,

$$P(v \in C \mid I_v, L_v) = \frac{P(I_v, L_v \mid v \in C)P(v \in C)}{P(I_v, L_v)}.$$

Since $P(I_v, L_v)$ is constant for each voxel across all classes, we obtain:

 $P(v \in C \mid I_v, L_v) \propto P(I_v, L_v \mid v \in C) P(v \in C).$

Assumption 2 L_v and I_v are conditionally independent of each other, given knowledge of the class to which v belongs:

 $P(I_{\nu}, L_{\nu} \mid \nu \in C) = P(I_{\nu} \mid \nu \in C)P(L_{\nu} \mid \nu \in C),$ or, equivalently,

$$P(v \in C \mid I_v, L_v) \propto P(I_v \mid v \in C) P(L_v \mid v \in C) P(v \in C).$$
(2)

In other words, we assume that there is no association between signal intensity and location within a brain structure of interest; the distribution of signal intensity within a structure is stationary. Note that this assumption is one of *conditional*, rather than of *marginal*, independence; if we do not know the classification of a voxel, its intensity and location are *not* independent of each other. In our application, by choosing classes or structures that have stationary signal-intensity distributions, we ensure that this assumption is valid.

Assumption 3 The signal-intensity noise in the image is white, additive, Gaussian, and tissue-independent. Under this commonly made assumption, the signal-intensity distribution of each class assumes a multivariate-normal distribution:

$$P(I_{v} | v \in C) = \frac{1}{\sqrt{(2\pi)^{d} |\Sigma_{c}|}} \exp\left(\frac{(I_{v} - \mu_{c})^{t} \Sigma_{c}^{-1} (I_{v} - \mu_{c})}{-2}\right)$$

where μ_{c} and Σ_{c} are the mean vector and covariance matrix for class *C*, respectively.

To construct a training set, we manually delineated lesions, then registered the images to a standard coordinate space. Although our algorithm could handle a much larger list of classes, for the purposes of the ACCORD-MIND study we consider four classes for classification: GM, WM, CSF, and lesions.

Our algorithm used the registered, segmented training data to compute the fractions of voxels in the training set that belonged to each class, to estimate P(C). In addition, our algorithm computed μ and Σ for each class to estimate the signal-intensity distribution for each class. Finally, our algorithm computed $P(L_v | v \in C)$ the fraction of all training-set voxels in class *C* that were found at location v.

Using only the signal-intensity, spatial, and prior distributions for each class, our algorithm can apply Equations 1 and 2 to each voxel v, labeling it as a lesion if the lesion class is most likely. Note that segmented lesions are localized in the standard coordinate space D; if necessary, we can use the inverse of a subject's spatial-registration transformation to register each segmented lesion back into that subject's original coordinate space.

Fig. 1 provides an overview of the training and lesionsegmentation processes. Before training or segmentation can occur, we perform a series of image-preprocessing steps, described in the next section; only skull stripping is not fully automated. In the training process, we compute three distributions for each class: the prior, P(C); the spatial model, $P(L_v | v \in C)$; and the signal-intensity distribution parameters, μ_c and Σ_c . In the lesion-segmentation process, new images are preprocessed, and are then classified based on Equations 1 and 2.

Image Preprocessing

Before we can use a subject's images for training or lesion detection, we must perform five preprocessing steps: coregistration of sequences; skull stripping, to remove non-brain tissues; segmentation into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF); spatial normalization *Figure 1.* General workflow for training and lesion segmentation. Shaded rectangles represent initial or processed image data; white rounded rectangles represent processing steps.



to bring that subject's sequences into alignment with a spatial standard that is common to all subjects; and intensity normalization, to ameliorate the effects of different scanners. The result of these image-processing steps is a set of sequences normalized with respect to spatial and signal-intensity characteristics.

Co-registration

Because the MR sequences cannot be acquired simultaneously, inter-sequence motion correction is necessary. Toward this end, we employed a rigid transformation based on mutual information [18]; we call this step *co-registration*. After co-registration, all sequences acquired for a subject are registered to the T1 image space for that subject.

Skull stripping

We refer to the removal of non-brain tissue, such as skull and skin, from the image volume as *skull stripping*. We applied a deformable-model–based skull stripping algorithm [19] to the T1-weighted images. Since image volumes for the other MR sequences were co-registered to the T1 image space of that subject, we used the T1 image volume as a reference to skullstrip the other sequences.

Brain-tissue segmentation

Since the elastic-registration algorithm [16] that we used to transform each subject to the standard space (see next paragraph) requires a WM/GM/CSF segmented image as input, we applied a segmentation algorithm based on hidden Markov random fields [20] to each T1-weighted volume.

Spatial Normalization and Smoothing

To register corresponding anatomic structures across subjects, we transformed each subject to an atlas developed by Kabani [21]; we refer to this coordinate space as the MNI space. Although the registration algorithm that we selected has been shown to register brain images with high accuracy, there will always be at least minimal misregistration [17]. Smoothing has been shown to ameliorate the effects of registration error [22]; therefore, we applied a 9-mm Gaussian kernel as in [23] to smooth the spatial-probability maps $P(L_y | v \in C)$.

Intensity Normalization

Due to the variations in image-acquisition hardware, maintenance, and software, we cannot assume that a given sequence will always have the same contrast and brightness. Compensating for these differences is particularly important when image data are obtained at different sites. Toward this end, we implemented intensity normalization by matching the histograms of each subject's images to those of a selected target subject [24]. As the intensity information for the ACCORD-MIND study consists of four MR sequences, we normalized each MR sequence separately.

Validation Methods

We validated our approach using leave-one-out cross validation (LOOCV): we performed 42 experiments, in each of which we used manually segmented MR images from 41 subjects to train our method, and the images from the remaining subject to test segmentation. Comparing automatically segmented lesions to the ground truth for the test subject, we computed the true-positive rate (*TPR*) and false-positive rate (*FPR*) of our approach:

$$TPR = \frac{TP}{TP + FN}$$
, and $FPR = \frac{FP}{FP + TN}$,

where *TP* and *FP* are the numbers of true- and false-positive voxels, and *TN* and *FN* are the numbers of true- and false-negative voxels, respectively.

We evaluated our algorithm using the following measures:

 ROC Curve: Recall that the lesions segmented by our algorithm are in a standard coordinate space, whereas manually segmented lesions were in the original sequence (e.g., FLAIR) image space for each subject. To make these results comparable, we transformed the automatically segmented lesions from the standard coordinate space back into the original subject's image space, applying a threshold to determine which voxels contained lesions after transformation. By varying this threshold, we obtained an ROC curve. **Similarity Index**: The similarity index (*SI*) is a commonly used metric of segmentation similarity [7]:

$$SI = \frac{2IP}{2TP + FP + FN}$$

Similarity-index computation is sensitive not only to total segmented-lesion volume, but also to lesion overlap [7]. We computed a *SI* curve for the LOOCV experiment by changing the threshold as we did when computing the ROC curve.

Because our data were obtained from two sites, using MR scanners from different manufacturers, it is important to determine the effects of these differences. We designed three experiments (*T1S1*, *T2S1-N*, and *T2S1*), based on three different training groups and one test group:

TIS1: We randomly selected 12 of the 30 subjects from site 1 to train the classifier, and we tested the classifier using the remaining 18 subjects from that side.

T2S1-N: We trained a classifier with all 12 subjects from site 2, and tested segmentation with the same 18 subjects used to test the *T1S1* classifier.

T2S1: We repeated the *T2S1-N* experiment, omitting the intensity-normalization step.

RESULTS

The scatter-gram for manually segmented lesions (ground truth) is shown in *Fig. 2*. The average lesion burden is 2.05 mL, and the standard deviation is 3.25 mL.

Fig. 3 shows the ROC curve for the LOOCV experiment; the area under the ROC curve (AUC) is 0.96. The *SI* curve is shown in *Fig. 4*; the maximum *SI* is 0.596, and the corresponding threshold is 0.75. *Fig. 5* shows the *SI* for each subject, computed using a threshold of 0.75. Note that some subjects have much higher *SI* values than others; as described previously [11,15], larger total lesion volumes tend to result in larger *SI* values, as seen in *Fig. 6*, which shows the relationship between *SI* and the relative lesion volume of each subject, which we define as the fraction of total brain volume occupied by manually segmented lesions.

The ROC curves for the multi-site experiments are shown in *Fig.* 7. Note that the T1S1 and T2S1-N curves are almost indistinguishable from each other, yet are readily distinguished from the T2S1 curve.

To provide a visual sense of classification results of the LOOCV experiments, we display representative images containing segmented lesions, and their corresponding manual-segmentation results of subjects with marked, moderate and minimal lesion burdens, in *Fig. 8, Fig. 9* and *Fig. 10*, respectively. *Fig. 11* shows rendered voxelwise lesion-prevalence rates across all subjects for automatically and manually segmented lesions.

We implemented our algorithm using the C++ programming language, on a Silicon Graphics (Mountain View, CA) Origin 300 workstation. Image preprocessing required approximately

Figure 2. Scatter-gram for manually segmented lesions.



Figure 4. Similarity index for LOOCV.



Figure 6. Similarity index versus relative lesion volume for each subject.



Figure 3. ROC curve for LOOCV.



Figure 5. Similarity index for each subject in LOOCV.



Figure 7. ROC curves for cross-site segmentation experiments. "T1" or "T2" indicates trained on data from site 1 or site 2; "S1" or "S2" indicates tested on data from site 1 or site2; N indicates intensity normalization.



Figure 8. Automatic versus manual lesion segmentation in a subject with large lesion burden. (a) T1; (b) T2; (c) SD; (d) FLAIR; (e) automatic segmentation; (f) manual segmentation.

Figure 9. Automatic versus manual lesion segmentation in a subject with moderate lesion burden. (a) T1; (b) T2; (c) SD; (d) FLAIR; (e) automatic segmentation; (f) manual segmentation.



260 minutes for each subject, of which the elastic-registration step required approximately 180 minutes. Our algorithm required approximately 16 minutes to compute a Bayesian classification model from 41 training subjects. The modelbuilding process required approximately 30 seconds for each additional training subject. Lesion detection for a new subject, not including preprocessing, required approximately 4 minutes.

DISCUSSION

Supervised versus unsupervised algorithms

Most segmentation algorithms belong to one of two categories: supervised [10,15] or unsupervised classification [13,25]. Although unsupervised methods do not require manually segmented lesions as training data, these methods usually require more user-specified parameters than do supervised methods.

Our Bayesian approach, being based on training data, and using Bayes' theorem to fuse prevalence, signal-intensity, and spatial information, requires minimal user input. For lesion segmentation, our method requires no user input; only image preprocessing-in particular skull stripping for the T1weighted sequence-requires some user guidance. Because the transformation of detected lesions from the standard space to a subject's space results in fractional rather than binary voxel values, the user may set a threshold for labeling transformed voxels as lesions, although the algorithm can apply the threshold that maximizes SI for training subjects. Because of the difficulties of brain-lesion segmentation, many methods [6,11,12] require user input, which precludes these methods' scaling to hundreds or thousands of subjects. A fully automatic classification method is critical for studies involving hundreds, or thousands, of subjects, such as ACCORD-MIND.

Figure 10. Automatic versus manual lesion delineation in a subject with minimal lesion burden. (a) T1; (b) T2; (c) SD; (d) FLAIR; (e) automatic segmentation; (f) manual segmentation.



f

Lesion class versus outlier detection

For brain-lesion-segmentation algorithms that are based on outlier detection [13,25,26], it is not necessary to model lesion characteristics explicitly; however, not all outliers are lesions [13], which will increase *FPR*. To remove outliers that are not lesions, additional constraints must be applied, which may differ across subjects. In contrast, we modeled lesions' statistical characteristics based on training data. If necessary, it would be straightforward to remove the lesion class from our algorithm, and to detect outliers using a Bayesian approach, for example, by computing the joint probability of the signalintensity and spatial data [27,28].

Relationship to other work

SI is one measure of how well automatic segmentation matches ground truth. However, larger lesion burdens increase *SI* [14,15], since a small volume of misclassification may constitute a large percentage of lesions when the lesion burden is low. This property renders *SI* values obtained from different data sets difficult to compare, unless those experiments share a common lesion prevalence. For example, the average volume of manually segmented lesions for our data set is 2.05 mL, compared to 15.0 mL for Anbeek's multiple-sclerosis segmentation research; it is not surprising that our maximal *SI*, 0.596, is smaller than the *SI* reported by Anbeek [15].

Although manually segmented lesions are often used as the ground truth to evaluate computer-aided algorithms [13-15], the notion of ground truth is elusive. Manual and automatic algorithms may differ because neither the expert nor the automatic method is perfectly valid or reliable. In general, our algorithm was more conservative in labeling lesions than the neuroradiologist. Based on visual inspection of the MR images and manual and computer-based segmentations, we suspect that differences between the two are primarily due to false-positive voxels during manual segmentation, rather than false-negative voxels during computer-based segmentation;

Figure 11. Lesion-distribution map resulting from the summation of cross-validation classification of all 42 subjects. The first row is the map of automatically detected lesions; the second row is the map of manually segmented lesions. Green represents ventricles; red represents lesions; white represents cortex.



however, there is no definitive way to prove this conjecture in the context of these experiments.

The spatial model that we used may require many training subjects if each subject has a low lesion burden. Although we designed our approach to support large-scale clinical research, we also accommodated small sample size by implementing an uninformative prior for the spatial model [29]. This initialization provides a reasonable distribution when few data are available for computing voxel-wise lesion prevalence, yet is discounted rapidly as sample size increases, yielding values virtually identical to observed voxel-wise prevalence when sample size is adequate.

Although our approach requires several imagepreprocessing steps, by separating these steps from Bayesian analysis, we greatly simplify comparison of competing registration, skull-stripping, and other image-processing methods, and can readily modify our software to employ those methods that we find to perform best for a particular study.

CONCLUSIONS

We have described the design, implementation, and evaluation of a Bayesian lesion-segmentation algorithm for brain images that synthesizes prior probabilities of class membership, spatial information, and multi-channel signal-intensity information. We used training subjects to obtain statistics for each segmentation class; we then used Bayesian methods to combine these distributions to classify images for new subjects. Experimental results using data from an ongoing clinical trial demonstrate that our approach has high accuracy, is robust to scanner and sequence-implementation differences, runs efficiently, and requires no user input for lesion segmentation.

ACKNOWLEDGEMENTS

The work was supported by the National Institutes of Health (NIH) under grant R01 AG13743, which is funded by the National Institute of Aging, the National Institute of Mental health, and the National Cancer Institute. This work was also supported by contracts N01-HC95178, N01-HC-95179, N01-HC-95180, N01-HC-95181, N01-HC-95182, N01-HC-95183, N01-HC-95184, IAA Y1-HC-9035, and IAA Y1-HC-1010 from the National Heart, Lung, and Blood Institute, the National Institute of Diabetes and Digestive and Kidney Diseases, the National Eye Institute, the National Institute on Aging, and the Centers for Disease Control and Prevention. This work was also supported by grant NHLBI- HC-99-16, which is funded by the NIH's General Clinical Research Centers, and by the Intramural Research Program at the National Institute on Aging.

REFERENCES

1. Breteler MM, van Swieten JC, Bots ML, Grobbee DE, Claus JJ, van den Hout JH, van Harskamp F, Tanghe HL, de Jong PT, van Gijn J. Cerebral white matter lesions, vascular risk factors, and cognitive function in a population-based study: the Rotterdam Study. Neurology. 1994 Jul;44(7):1246-52.

2. Haan MN, Shemanski L, Jagust WJ, Manolio TA, Kuller L. The role of APOE e4 in modulating effects of other risk factors for cognitive decline in elderly men. JAMA. 1999 Jul 7;282(1):40-6.

3. de Leeuw FE, de Groot JC, Oudkerk M, Witteman JC, Hofman A, van Gijn J, Breteler MM. Hypertension and cerebral white matter lesions in a prospective cohort study. Brain. 2002 Apr;125(Pt 4):765-72.

 de Groot JC, de Leeuw FE, Oudkerk M, Hofman A, Jolles J, Breteler MM. Cerebral white matter lesions and depressive symptoms in elderly adults. Arch Gen Psychiatry. 2000 Nov;57(11):1071-6.

5. Kuller LH, Shemanski L, Manolio T, Haan M, Fried L, Bryan N, Burke GL, Tracy R, Bhadelia R. Relationship between ApoE, MRI findings, and cognitive function in the Cardiovascular Health Study. Stroke. 1998 Feb;29(2):388-98.

6. Udupa JK, Wei L, Samarasekera S, Miki Y, van Buchem MA, Grossman RI. Multiple sclerosis lesion quantification using fuzzy-connectedness principles. IEEE Trans Medical Imaging. 1997 Oct;16(5):598-609.

7. Zijdenbos AP, Dawant BM, Margolin RA, Palmer AC. Morphometric analysis of white matter lesions in MR images: method and validation. IEEE Trans Med Imaging. 1994;13(4):716-24.

8. Williamson JD, Miller ME, Bryan RN, Lazar RM, Coker LH, Johnson J, Cukierman T, Horowitz KR, Murray A, Launer LJ; ACCORD Study Group. The action to control cardiovascular risk in diabetes memory in diabetes study (ACCORD-MIND): rationale, design, and methods. Am J Cardiol. 2007 Jun 18;99(12A):112i-122i.

9. Pachai C, Zhu YM, Grimaud J, Hermier M, Dromigny-Badin A, Boudraa A, Gimenez G, Confavreux C, Froment JC. A pyramidal approach for automatic segmentation of multiple sclerosis lesions in brain MRI. Comput Med Imaging Graph. 1998 Sep-Oct;22(5):399-408.

10. Kamber M, Shinghal R, Collins DL, Francis GS, Evans AC. Model-based 3-D segmentation of multiple sclerosis lesions in magnetic resonance brain images. IEEE Trans Med Imaging. 1995;14(3):442-53.

11. Johnston B, Atkins MS, Mackiewich B, Anderson M. Segmentation of multiple sclerosis lesions in intensity corrected multispectral MRI. IEEE Trans Med Imaging. 1996;15(2):154-69.

12. Hojjatoleslami SA, Kruggel F. Segmentation of large brain lesions. IEEE Trans Med Imaging. 2001 Jul;20(7):666-9.

13. van Leemput K, Maes F, Vandermeulen D, Colchester A, Suetens P. Automated segmentation of multiple sclerosis lesions by model outlier detection. IEEE Trans Med

Imaging. 2001 Aug;20(8):677-88.

14. Admiraal-Behloul F, van den Heuvel DMJ, Olofsen H, van der Grond J, van Buchem MA, Reiber JH. Fully automatic segmentation of white matter hyperintensities in MR images of the elderly. Neuroimage. 2005 Nov 15;28(3):607-17.

15. Anbeek P, Vincken KL, van Osch MJP, Bisschops RH, van der Grond J. Automatic segmentation of differentsized white matter lesions by voxel probability estimation. Med Image Anal. 2004 Sep;8(3):205-15.

16. Shen D, Davatzikos C. HAMMER: Hierarchical attribute matching mechanism for elastic registration. IEEE Trans Med Imaging. 2002 Nov;21(11):1421-39.

17. Shen D, Davatzikos C. Very high-resolution morphometry using mass-preserving deformations and HAMMER elastic registration. Neuroimage, 2003. 18(1): p. 28–41. 2003 Jan;18(1):28-41.

18. Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. IEEE Trans Med Imaging. 1997 Apr;16(2):187-98.

19. Smith SM. Fast robust automated brain extraction. Hum Brain Mapp. 2002 Nov;17(3):143-55.

20. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans Med Imaging. 2001 Jan;20(1):45-57.

21. Kabani N, MacDonald D, Holmes C, et al. A 3D atlas of the human brain. Proceedings of the Human Brain Mapping Conference; 1998; Montreal. Abstract 717.

22. Shen D, Liu D, Liu H, Clasen L, Giedd J, Davatzikos C. Automated morphometric study of brain variation in XXY males. Neuroimage. 2004 Oct;23(2):648-53.

23. Chen R, Herskovits EH. Graphical-model–based morphometric analysis. IEEE Trans Med Imaging. 2005 Oct;24(10):1237-48.

24. Freedman D, Radke RJ, Zhang T, Jeong Y, Lovelock DM, Chen GT. Model-based segmentation of medical imagery by matching distributions. IEEE Trans Med Imaging. 2005 Mar;24(3):281-92.

25. Prastawa M, Bullitt E, Ho S, Gerig G. A brain tumor segmentation framework based on outlier detection. Med Image Anal. 2004 Sep;8(3):275-83.

26. Aït-Ali LS, Prima S, Hellier P, Carsin B, Edan G, Barillot C. STREM: A robust multidimensional parametric method to segment MS lesions in MRI. Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv. 2005;8(Pt 1):409-16.

27. Roberts SJ. Extreme value statistics for novelty detection in biomedical data processing. IEE Proceedings of Science, Measurement & Technology, 2000. 147(6): p. 363–367.

28. Tarter ME. Density estimation applications for outlier detection. Comput Programs Biomed. 1979 Jul;10(1):55-60.

29. Berger JO. Statistical Decision Theory and Bayesian Analysis. New York: Springer-Verlag; 1985.