AJNR

Automatic Lesion Incidence Estimation and Detection in Multiple Sclerosis Using Multisequence Longitudinal MRI

E.M. Sweeney, R.T. Shinohara, C.D. Shea, D.S. Reich and C.M. Crainiceanu

Automatic Lesion Incidence Estimation and Detection in Multiple Sclerosis Using Multisequence Longitudinal MRI

E.M. Sweeney, R.T. Shinohara, C.D. Shea, D.S. Reich, and C.M. Crainiceanu

ABSTRACT

BACKGROUND AND PURPOSE: Detecting incidence and enlargement of lesions is essential in monitoring the progression of MS. In clinical trials, lesion load is observed by manually segmenting and comparing serial MR images, which is time consuming, costly, and prone to inter- and intraobserver variability. Subtracting images from consecutive time points nulls stable lesions, leaving only new lesion activity. We propose SuBLIME, an automated method for segmenting incident lesion voxels.

MATERIALS AND METHODS: We used logistic regression models incorporating multiple MR imaging sequences and subtraction images from consecutive longitudinal studies to estimate voxel-level probabilities of lesion incidence. We used T1-weighted, T2-weighted, FLAIR, and PD volumes from a total of 110 MR imaging studies from 10 subjects.

RESULTS: To assess the performance of the model, we assigned 5 subjects to a training set and the remaining 5 to a validation set. With SuBLIME, lesion incidence is detected and delineated in the validation set with an AUC of 99% (95% CI [97%, 100%]) at the voxel level.

CONCLUSIONS: This fully automated and computationally fast method allows sensitive and specific detection of lesion incidence that can be applied to large collections of images. Using the explicit form of the statistical model, SuBLIME can easily be adapted to cases when more or fewer imaging sequences are available.

ABBREVIATIONS: AUC = area under ROC curve; IQR = interquartile range; NAWM = normal appearing white matter; PD = proton attenuation; ROC = receiver operating characteristic; SuBLIME = Subtraction-Based Logistic Inference for Modeling and Estimation

MS is an inflammatory disease of the central nervous system characterized by brain and spinal cord lesions. Although lesions in the gray matter of the brain are common, lesions are more readily recognized in the white matter.[1] MR imaging of the brain is used to detect lesions in MS and is essential in monitoring disease progression. In addition to documenting disease effects at one time, MR imaging can be used to assess longitudinal changes.[2] MR imaging observation of lesion volume change resulting from the development of new lesions, enlarging lesions, and resolving lesions is an important marker of disease progression and response to therapy.[3] Lesion volume change is a common outcome in clinical trials and is computed by comparing manual segmentations of serial MR imaging examinations,[4] which is time-consuming, costly, and subject to inter- and intraobserver variability. In addition, quantifying lesion change is challenging because incident and enlarging lesions represent only a small proportion of all lesions, typically on the order of 5–10%.[5] Therefore, a method that can automatically and accurately segment lesions to assess change is desirable.

Lladó et al[6] provide a comprehensive review of the current methods of lesion change segmentation by dividing them into 2 groups: intensity-based and deformation-based. Intensity-based methods use voxel-level comparisons of intensities to identify lesion change. These methods are further categorized into deterministic, statistical, and temporal analysis methods. Deformation-based methods rely on observing the changes in tissue surrounding lesions due to lesion expansion or contraction.

Deterministic methods are defined as segmentations of sub-

Please address correspondence to Elizabeth Sweeney, Department of Biostatistics, Johns Hopkins University, 615 N Wolfe St, Baltimore, MD 21205; e-mail:emsweene@jhsph.edu

traction images.[6] Images from consecutive studies can be registered, normalized, and then subtracted to isolate areas of lesion change and cancel radiologically stable disease-related measurements. Manually segmented 2D T2-weighted subtraction images identify a higher number of active lesions, with greater inter- and intraobserver agreement[5,7] than comparing independently segmented serial images. Subtraction images created from 3D imaging acquisitions have been shown to be less susceptible to artifacts than 2D subtraction images.[8] To our knowledge, no manual or automated method for combining information from multiple types of MR imaging subtraction images has been developed.

Current statistical methods detect changes in the brain between 2 MR imaging studies but do not provide fully automated segmentations of lesion change.[6] Volumes from 2 time points are registered, and statistical models classify voxels as no signal change, signal increase, signal decrease, and outside of the brain.[9] Higher lesion activity detection probabilities have been obtained by extending this method to the use of multisequence imaging data,[10] but clusters of highly likely lesion change voxels must be reviewed by an expert to produce lesion change segmentations.

Temporal analysis approaches are based upon voxel level time-series.[6] Longitudinal collections of images for a subject are registered, and voxel intensities are analyzed as a function of time. Lesions and other tissue classes of the brain can be segmented from these time-series,[11,12] but these methods require more than 2 MR imaging studies for each subject.

In contrast to intensity-based methods, deformation field-based approaches observe the modification of the tissue surrounding lesions.[6] Volumes from 2 time points are registered, and structural changes in the brain are determined by the local deformation of voxels.[13,14] While these methods have been shown to perform better than intensity-based methods for the segmentation of lesions, they require a region of interest to be manually selected before the analysis.

There is currently no fully automated method that produces segmentations of incident lesions and lesion enlargement using information from only 2 MR imaging studies. We therefore propose SuBLIME, a fully automated and computationally fast method for segmenting voxel-level lesion incidence between 2 MR imaging studies. Using a logistic regression model with various MR imaging sequences from consecutive studies, we estimate the probability of lesion incidence in each voxel. The model utilizes information from the FLAIR, PD, T2-weighted, and T1-weighted volumes, as well as corresponding subtraction images, to produce interpretable results in the form of regression coefficients that can be applied to new pairs of images quickly and easily.

To show how multisequence image information improves lesion incidence and enlargement detection, while also being robust to artifacts, we compare the full SuBLIME model to a nested model fit with only the T2-weighted image. We also compare both models to a population-level threshold of the T2-weighted subtraction image to show that the SuBLIME model framework improves performance.

Table 1: Subject demographic, diagnosis, and treatment information

Set	Subtype	Age	Sex	EDSS	Treatment with Disease-Modifying Therapy (Baseline)
Validation	RRMS	37	Male	1.5	Yes
Validation	RRMS	38	Female	2.0	Yes
Validation	RRMS	48	Male	3.0	No
Training	RRMS	38	Female	1.5	No
Training	RRMS	30	Female	1.0	Yes
Training	RRMS	43	Female	1.5	Yes
Validation	RRMS	35	Female	1.5	Yes
Training	RRMS	37	Female	1.0	Yes
Training	RRMS	47	Female	3.0	Yes
Validation	RRMS	56	Female	1.0	No

Note:—EDSS indicates Expanded Disability Status Scale; RRMS, relapsing-remitting multiple sclerosis.

Table 2: Study scanning parameters

	FA (degrees)	TR (ms)	TE (ms)	TI (ms)
FLAIR	(90, 90)	(10,000, 10,000)	(77.8, 159.5)	(2200, 2500)
T2-weighted	(90, 90)	(3400, 6500)	(94.6, 112.0)	NA
PD	(90, 90)	(3400, 6500)	(11.8, 15.0)	NA
T1-weighted	(13, 20)	(7.68, 10.3)	(1.88, 4.05)	(450, 750)[a]

Note:—FA indicates flip angle; NA, not available.
[a] 106 of the T1-weighted scans did not have an inversion time.

MATERIALS AND METHODS

Study Population

A total of 110 MR imaging studies (11 longitudinal studies each of 10 subjects) were analyzed. All participants gave written consent and were scanned as part of an institutional review board–approved natural history protocol. To assess the performance of the model, we randomly assigned 5 subjects (55 studies) to a set for model training and used the remaining 5 subjects (55 studies) for validation of our model. Demographic, diagnosis, and treatment information for the training and validation subjects can be found in Table 1. The mean time between scans in the training set was 130 days and the median was 35 days. The mean time between scans in the validation set was 88 days and the median was 34 days.

Experimental Methods

We acquired whole-brain 2D FLAIR, PD, T2-weighted, and 3D T1-weighted volumes in a 1.5T MR imaging scanner (Signa Excite HDxt; GE Healthcare, Milwaukee, Wisconsin) using the body coil for transmission. For signal detection, we used a volume head coil for 107 of the studies and an 8-channel receive coil (Invivo, Gainesville, Florida) in the remaining 3 studies. The 2D FLAIR, PD, and T2-weighted volumes were acquired using spin-echo sequences and the 3D T1-weighted volume using a gradient-echo sequence. The PD and T2-weighted volumes were acquired as short and long echoes from the same sequence. We used various scanning parameters for the different studies. The ranges for the flip angle, repetition time, echo time, and inversion time can be found in Table 2.

Image Postprocessing

For initial image processing, we used Medical Image Processing Analysis and Visualization (http://mipav.cit.nih.gov) and Java Image Science Toolkit (http://www.nitrc.org/projects/jist/).[15] We interpolated all images for each subject to a voxel size of 1 mm³

and rigidly aligned all of the T1-weighted volume to the Montreal Neurological Institute standard space. We then took the average of the T1-weighted volume for all the studies for each subject and registered the longitudinal collection of the T1-weighted, T2-weighted, FLAIR, and PD images for the subject to this average image. We removed extracerebral voxels using a skull-stripping procedure.[16] We automatically segmented the entire brain and the NAWM using the T1-weighted and FLAIR images.[17]

Criterion Standard Measure

To fit the model and to measure performance, we required a set of data in which the outcome is assessed by using a criterion standard measure. The criterion standard was obtained using manual segmentation by an experienced neuroradiologist, who evaluated consecutive T2-weighted and FLAIR images and the T2-weighted subtraction image. For each study, we segmented 5 representative axial sections of the brain: sections 50 (through the inferior temporal lobes, pons, and cerebellum), 70 (including the midbrain and occipital lobes), 90 (through the internal capsules), 110 (through the centrum semiovale), and 130 (near the vertex). Only voxels with new lesion activity were segmented; voxels containing existing lesions were not segmented. In total, there were 500 segmented sections, with 23 sections containing lesions. There were a total of 55 incident or enlarging lesions, with 26 in the training set and 29 in the validation set. The neuroradiologist only segmented changes between consecutive studies; in our analysis, we also only focus on incidence between consecutive studies.

Normalization

We performed all statistical modeling in the R environment (version 2.12.0; R Foundation for Statistical Computing, Vienna, Austria) with the packages AnalyzeFMRI,[18] biglm,[19] ff,[20] and ROCR.[21] We used intensities from the FLAIR, PD, T2-weighted, and T1-weighted volumes to identify new lesions and enlargement of previously existing lesions. We denote the observed intensity of voxel v, for subject i, from MR imaging study conducted at time t_j by

$$M_i^0(v, t_j), \ M = FLAIR, \ PD, \ T2, \ T1$$

where M indicates the imaging sequence.

Analyzing images across study visits requires that the images be normalized so that voxel intensities have common interpretations. The normalization must allow for comparison of studies within and between subjects. Our aim was to segment incident lesions, which occur predominantly in the white matter; therefore, intensities from each study are expressed as a departure from the subject's NAWM mean in each imaging sequence[22]:

$$M_i^N(v, t_j) = \{M_i^0(v, t_j) - \mu_{i,M}^0(t_j)\}/\sigma_{i,M}^0(t_j)$$

where $\mu_{i,M}^0(t_j)$ and $\sigma_{i,M}^0(t_j)$ are the mean and standard deviation of the observed voxel intensities in the NAWM of subject i, from sequence M, conducted at time t_j. The NAWM segmentations were developed using the high-resolution T1-weighted and FLAIR images.[17] SuBLIME can use any NAWM segmentation approach and is not sensitive to the choice of segmentation algorithm.

Subtraction Images

We calculated subtraction images for each sequence by subtracting normalized images from consecutive MR imaging studies. We denote the subtraction image intensity at voxel v for study t_j by

$$\Delta M_i^N(v, t_j) = M_i^N(v, t_j) - M_i^N(v, t_{j-1})$$

where i is the subject and M is the imaging sequence. Misregistration, partial volume effects, and low signal-to-noise ratios can cause artifacts and noise in subtraction images[2]; our approach avoids these pitfalls by carefully integrating population-level normalization with robustly tuned statistical prediction algorithms.

Subtraction-Based Logistic Inference for Modeling and Estimation

SuBLIME uses logistic regression to model the probability that a voxel is part of an incident or enlarging lesion. We model lesion incidence at the voxel level using FLAIR, PD, T2-weighted, and T1-weighted intensities from the current study; the subtraction image intensities for each of these sequences; the time between the consecutive studies; and the interactions between the subtraction image intensities and the time between consecutive studies. The result of our model is a collection of coefficients that can be used to create 3D maps of the probabilities of lesion incidence. A flow chart describing the SuBLIME method can be found in the On-line Appendix.

The first step of this procedure is to isolate candidate voxels that are likely to be part of incident or enlarging lesions. As changes in lesion behavior are evident on T2-weighted subtraction images, we apply the logistic regression model only to voxels that are hyperintense on these images. To avoid isolated voxels that have high T2-weighted subtraction values as a result of noise, we smooth the subtraction image by using a Gaussian kernel with a window size of 5 mm. Then, we include only voxels that have T2-weighted subtraction values larger than 1 standard deviation of the subtraction intensities calculated across the image. We chose the threshold of 1 standard deviation on the T2-weighted subtraction image empirically and found it to be liberal in capturing changes in lesions. Of the 1026 neuroradiologist-segmented lesion voxels in the training set, the voxel selection mask excluded only 45, or 4%. At least 1 voxel from every lesion was included in the candidate voxels. The 45 missing voxels were found in the voxel selection masks for 4 studies from 3 of the subjects in the training set. The missed voxels were evenly distributed among the sections with lesion incidence. On average, for each pair of consecutive studies in the training set, 14,311 (IQR: 9965; 17,638) voxels were included as candidate voxels in the 5 representative sections of the brain. The average number of voxels in the brain from the 5 sections of the subjects in the training set is 59,116 (IQR: 56,269; 61,238). As the voxel selection procedure only excludes 4% of active lesion voxels, it has a negligible impact on sensitivity, but it greatly improves specificity by eliminating more than 75% of the candidate voxels for active lesions. The voxel selection mask is illustrated in Fig 1; Fig 1A shows the radiologist segmentation of an axial section of the brain, and Fig 1B shows the candidate voxels for the same section that will be used in the logistic modeling described here.

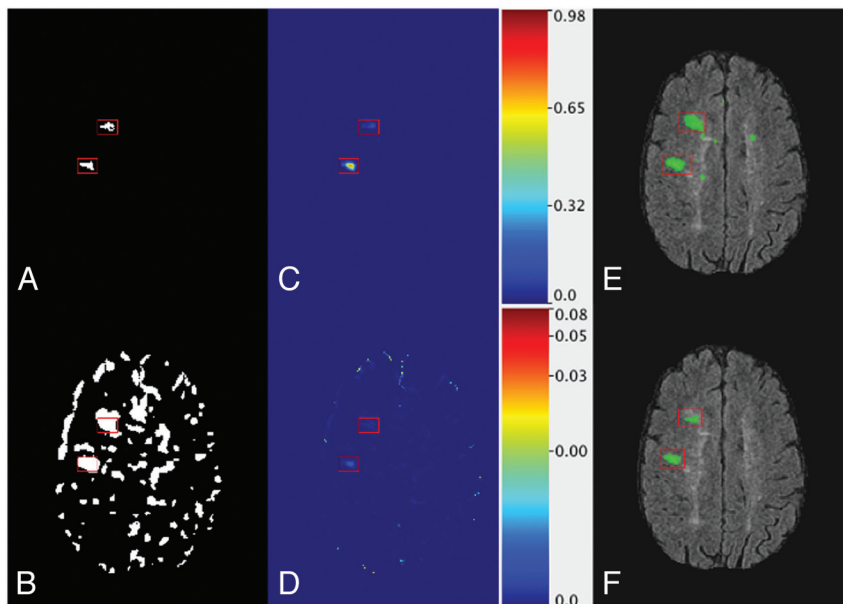We then fit a voxel-level logistic regression model for lesion

FIG 1. Areas with lesion incidence are indicated with *red boxes*. *A*, Neuroradiologist manual segmentation of an axial section of the brain. *B*, Selected voxels for SuBLIME modeling. *C*, Axial section of the probability map from the full model. *D*, Axial section of the probability map from the SuBLIME model fit with only the T2-weighted image. *E*, Binary segmentation of the probability map from the full model with false-positive rate of 0.01 overlaid on the FLAIR image. *F*, Binary segmentation of the probability map from the full model with false-positive rate of 0.001 overlaid on the FLAIR image.

incidence over these candidate voxels, denoting the voxel-level lesion incidence by the random variable *W*:

$$W_i(v, t_j)$$

$$= \begin{cases} 1, & \text{if subject } i \text{ has lesion incidence in voxel } v \text{ at time } t_j \\ 0, & \text{otherwise} \end{cases}$$

We model the probability that a voxel is part of a lesion incidence using the logistic regression model:

1) $\text{logit}\,[P\{W_i(v, t_j) = 1\}] = \beta_0 + \beta_1 \Delta t +$

$\beta_2 FLAIR_i^N(v, t_j) + \beta_3 \Delta FLAIR_i^N(v, t_j) + \beta_4 \Delta FLAIR_i^N(v, t_j) \times \Delta t +$
$\beta_5 PD_i^N(v, t_j) + \beta_6 \Delta PD_i^N(v, t_j) + \beta_7 \Delta PD_i^N(v, t_j) \times \Delta t +$
$\beta_8 T2_i^N(v, t_j) + \beta_9 \Delta T2_i^N(v, t_j) + \beta_{10} \Delta T2_i^N(v, t_j) \times \Delta t +$
$\beta_{11} T1_i^N(v, t_j) + \beta_{12} \Delta T1_i^N(v, t_j) + \beta_{13} \Delta T1_i^N(v, t_j) \times \Delta t$

where Δt is the time in days between consecutive studies. Note, however, that SuBLIME is a general logistic procedure encompassing a large number of logistic regression models. Here we use the specific model (1) because it performs very well in our application. The estimated coefficients from the full model and the nested model with only the T2-weighted image, based on all of the 10 subjects, are given in the On-line Appendix. To account for the correlation in patients over time, we nonparametrically bootstrapped (with replacement) the subjects for the training and validation set to allow for estimation of the confidence intervals for the regression coefficients. An interpretation of the coefficients is also provided in the On-line Appendix.

After we fit the model, we use the estimated coefficients to create maps of the estimated probability of lesion incidence at each voxel. To incorporate spatial information of the neighboring voxels and reduce noise, we smoothed the estimated probabilities from the model using a Gaussian kernel with window size of 3 mm. This kernel size was empirically chosen and found to perform well.

RESULTS

Axial sections from normalized FLAIR, PD, T2-weighted, and T1-weighted volumes from 2 consecutive studies, and the subtraction images of the 2 studies for each sequence for a subject in the validation set, are shown in the On-line Appendix. Lesions are characterized by hyperintensities in FLAIR, PD, and T2-weighted images, and hypointensities in T1-weighted images, though not all lesions appear on T1-weighted images. Lesion incidence is characterized by hyperintensities in FLAIR, PD, and T2-weighted subtraction images and by hypointensities in T1-weighted subtraction images. For the same subject and study, an axial section of the smoothed probability map of lesion incidence from the full SuBLIME model is shown in Fig 1*C*. Fig 1*D* shows the same section for a model fit with only the T2 image. Red indicates areas with the higher probability of lesion incidence, and blue indicates areas with lower probabilities of lesion incidence. An axial section of the probability map from the full model for another subject from the validation set is shown in the On-line Appendix. The corresponding axial section of the normalized FLAIR, PD, T2-weighted, and T1-weighted volumes from the 2 consecutive studies and the subtraction images are also shown.

In the Study Population section, we discussed assessing the performance of the model by randomly assigning 5 subjects to the training set and the remaining 5 subjects to a validation set. The voxel-level ROC curves for the subjects in the validation set are shown in Fig 2*A*. We optimized the voxel selection procedure and fit the model (1) on the training set. We only used studies from the validation set to estimate the voxel-level ROC curve and AUC. The vertical axis of the ROC curve shows the true-positive rate (sensitivity) for a given threshold on the probability map, and the horizontal axis shows the false-positive rate (1 − specificity) for this threshold. These performance measures are known to be susceptible to instability. To account for this instability, as well as the correlation in subjects over time, we nonparametrically bootstrapped (without replacement) the training and validation sets. This procedure allowed the estimation of confidence intervals for the AUC. The blue curve is the ROC curve for the full model with an estimated AUC of 99% (95% CI [97%, 100%]). The red curve is the ROC curve for the model fit with only the T2-weighted image with an estimated AUC of 97% (95% CI [88%, 99%]). The green curve is the ROC curve for the population level threshold of the T2-weighted subtraction image with an estimated AUC of 92% (95% CI [83%, 95%]).
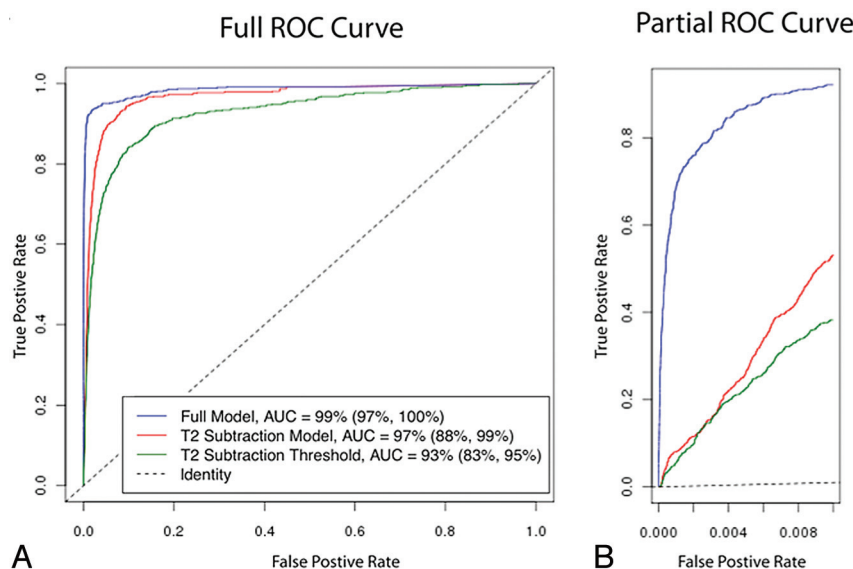
Full ROC Curve

Partial ROC Curve

FIG 2. *A*, ROC curves for the voxel-level detection of incident and enlarging lesions for different thresholds of the probability maps produced from SuBLIME, as well as different thresholds of the T2-weighted subtraction without the model and voxel selection procedure. The *red* ROC curve is for the full model and has an AUC of 99% (95% CI [97%, 100%]). The *blue* ROC curve is for the model fit with only the T2-weighted image and has an AUC of 97% (95% CI [88%, 99%]). The *green* ROC curve is for thresholding the T2-weighted subtraction image without the model and voxel selection procedure and has an AUC of 92% (95% CI [83%, 95%]). *B*, Partial ROC curves for false-positive rates up to 0.01. The *red* curve is for the full model and the *blue* curve is for the model fit with only the T2-weighted image. The *green* curve is for thresholding the T2-weighted subtraction image without the model and voxel selection procedure.

Table 3: Binary segmentations

False-Positive Rate	Threshold Value	Specificity	Sensitivity	Volume Change (Actual 625 mm³)
0.01	0.0022	0.99	0.95	30454
0.001	0.0229	0.999	0.83	3520
0.00025	0.0815	0.99975	0.54	1082
0.0001	0.1396	0.9999	0.35	509

The AUC calculated over the entire ROC space shown in Fig 2*A* is not optimal for the evaluation of the performance of an algorithm designed to detect lesion incidence. This full AUC summarizes test performance over regions of the ROC space that are not clinically relevant. The subset of lesions showing change usually consists of a small fraction of all lesions in the brain and an even smaller fraction of the entire brain. The average number of voxels in the entire brain among the validation set is 1,277,736 (IQR: 1,204,162; 1,364,908). For the entire brain, a false-positive rate of 0.01 would correspond to a volume of 12,800 mm³ of healthy brain being falsely identified as lesion, making the resulting prediction maps difficult to interpret and clinically unacceptable. Thus, the performance of SuBLIME should only be evaluated for very small false-positive rates; large false-positive rates are not clinically relevant. Fig 2*B* shows the partial ROC curve for false-positive rates up to 0.01. The blue curve corresponds to the full model (1), the red curve corresponds to the model fit with only the T2-subtraction image, and the green curve corresponds to the population level threshold of the T2-weighted subtraction image. The full model has a higher true-positive rate for these relevant small false-positive rates, which further emphasizes the importance of using the various imaging sequences in addition to the T2-weighted images. We have also provided the partial ROC curves for the nested SuBLIME models fit with only 3 imaging sequences in the On-line Appendix.

The probability maps can be cut at desired false-positive rates to create binary segmentations of lesion incidence. Table 3 shows the threshold values, specificity, and sensitivity for 4 different false-positive rates in the validation set. To put this into context, we have also provided the volume of lesion incidence that was segmented at each threshold value. The true lesion volume change from the neuroradiologist's manual segmentation was 625 mm³. Fig 1*E* shows a binary segmentation of a section for a false-positive rate of 0.01, and Fig 1*F* shows a binary segmentation with a false-positive rate of 0.001. The section with false-positive rate 0.01 contains many false-positive lesions, while the false-positive rate of 0.001 provides a more accurate segmentation of the section. Choosing threshold values is a trade-off between sensitivity and specificity. SuBLIME is flexible, and the appropriate false-positive rate may be selected for a particular application.

DISCUSSION

Lesion volume change is frequently an outcome in clinical trials for patients with MS and is traditionally assessed by manually segmenting consecutive MRI. In a clinical trial setting, SuBLIME may be used to replace or assist manual segmentation of incident and enlarging lesions to reduce costs. After training, our fully automatic method does not require human input and avoids the variability introduced by manual segmentation. Training is necessary for each new dataset, but can be fairly limited, as in the example from this paper.

In contrast to many automatic segmentation techniques, SuBLIME is computationally fast. Training the model on 5 sections from 110 images takes less than an hour on a standard workstation. This process is only conducted once after collecting a dataset that combines images from multiple protocols, and the results may be summarized as the 14 coefficients in the model (1). Using this fitted model to generate a probability map of the entire brain from a set of new images takes only seconds using a standard workstation. For this analysis, we trained and validated the model on 5 representative sections of the brain. Confirmatory analyses that train and validate on segmentations of the entire brain are indicated for further understanding of the performance of SuBLIME.

Because SuBLIME is fit voxelwise, it is sensitive to major misregistration within a study and between longitudinal studies for the same subject. However, SuBLIME is robust to minor errors in registration. By simultaneously comparing data from multiple se-

quences and only considering candidate voxels, SuBLIME distinguishes between artifacts and lesion incidence. We suspect SuBLIME may also be sensitive to the radiologist's manual segmentation. In this analysis, we used a single expert segmentation of the sections. Future studies will compare the SuBLIME segmentations to additional manual segmentations to investigate the influence of interrater variability.

SuBLIME uses a voxel-level model for assessing the outcome. The assumption of independence between voxels is imperfect, as lesions consist of clusters of voxels. In this work, we use smoothing of the T2-weighted subtraction image in candidate voxel selection, followed by a second smoothing of the predicted probabilities of the model to incorporate the spatial nature of the data. Nevertheless, further incorporation of neighboring voxel information is warranted.

For this analysis, we decided to use a group of patients with early relapsing disease because new lesion activity is more common in this group. In the future, we plan to apply SuBLIME to a larger group of patients with a wide spectrum of disease activity and severity. Although developed for the particular application of longitudinal lesion incidence segmentation in MS, SuBLIME can be applied more generally to coregistered serial images to detect other changes and pathologies. We expect that a version of SuBLIME can easily be adapted to monitor other pathologies assessed using MR imaging, such as volumetric changes in patients with vascular disease or tumors. Our techniques may also be useful for monitoring changes in other imaging outcomes, including studies that use combinations of imaging techniques such as PET-CT studies in oncology. Finally, it is likely that the methods can be applied to organs outside the brain, including the lung, liver, and kidneys.

CONCLUSIONS

SuBLIME is a fully automated and computationally fast method that allows sensitive and specific detection of lesion incidence and enlargement. Using the explicit form of the statistical model, SuBLIME can easily be adapted to cases when more or fewer imaging sequences are available.

REFERENCES

1. Sahraian AM, Radue EW. *MRI Atlas of MS Lesions*. Berlin, Germany: Springer-Verlag; 2007:178
2. Duan Y, Hildenbrand PG, Sampat MP, et al. **Segmentation of subtraction images for the measurement of lesion change in multiple sclerosis.** *AJNR Am J Neuroradiol* 2008;29:340–46
3. Filippi M, Rocca MA. **MR imaging of multiple sclerosis.** *Radiology* 2011;259:659–81
4. Simon JH, Li D, Traboulsee A, et al. **Standardized MR imaging protocol for multiple sclerosis: consortium of MS centers consensus guidelines.** *AJNR Am J Neuroradiol* 2006;27:455–61
5. Tan IL, Schijndel RAV, Miller DH, et al. **Image registration and subtraction to detect active T(2) lesions in MS: an interobserver study.** *J Neurol* 2002;249:767–73
6. Lladó X, Ganiler O, Oliver A, et al. **Automated detection of multiple sclerosis lesions in serial brain MRI.** *Neuroradiology* 2011 Dec 20 [Epub ahead of print]
7. Moraal B, Meier DS, Poppe PA, et al. **Subtraction MR images in a multiple sclerosis multicenter clinical trial setting.** *Radiology* 2009;250:506–14
8. Morall B, Wattjes MP, Geurts JG, et al. **Improved detection of active multiple sclerosis lesions: 3D subtraction imaging.** *Radiology* 2010;255:155–63
9. Lemieux L, Wieshmann U, Moran N, et al. **The detection and significance of subtle changes in mixed-signal brain lesions by serial MRI scan matching and spatial normalization.** *Med Image Anal* 1998;2:227–42
10. Bosc M, Heitz F, Armspach J, et al. **Automatic change detection in multimodal serial MRI: application to multiple sclerosis lesion evolution.** *NeuroImage* 2003;20:643–56
11. Gerig G, Martin J, Kikinis R, et al. **Unsupervised segmentation of 3-D dual-echo MR head data.** *Image Vis Comput* 1992;10:349–60
12. Welti D, Gerig G, Radü EW, et al. **Spatiotemporal segmentation of active multiple sclerosis lesions in serial MRI data.** *Inf Proc Med Imaging* 2001;2082:438–45
13. Thirion JP, Calmon G. **Deformation analysis to detect and quantify active lesions in three-dimensional medical image sequences.** *IEEE Trans Med Imaging* 1999;18:429–41
14. Rey D, Subsol G, Delingette H, et al. **Automatic detection and segmentation of evolving processes in 3D medical images: applications to multiple sclerosis.** *Med Image Anal* 2002;6:163–79
15. Lucas BC, Bogovic JA, Carass A, et al. **The Java Image Science Toolkit (JIST) for rapid prototyping and publishing of neuroimaging software.** *Neuroinformatics* 2010;8:5–17
16. Carass A, Wheeler MB, Cuzzocreo J, et al. **A joint registration and segmentation approach to skull stripping.** In *4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro* 2007;656–59
17. Shiee N, Bazin PL, Ozturk A, et al. **A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions.** *NeuroImage* 2010;49:1524–35
18. Bordier C, Dojat M, Lafaye De Micheaux P. *AnalyzeFMRI: an R package to perform statistical analysis on FMRI datasets*. R Package, version 1.1-12. 2009, http://CRAN.R-project.org/web/packages/AnalyzeFMRI/index.html
19. Lumley T. *biglm: bounded memory linear and generalized linear models*. R package, version 0.7. 2009. http://CRAN.R-project.org/package=biglm
20. Adler D, Glaeser C, Nenadic O, et al. *ff: memory-efficient storage of large data on disk and fast access functions*. R package, version 2.2-2. 2011. http://CRAN.R-project.org/package=ff
21. Sing T, Sander O, Beerenwinkel N, et al. *ROCR: visualizing the performance of scoring classifiers*. R package, version 1.0-4. 2009. http://CRAN.R-project.org/package=ROCR
22. Shinohara RT, Crainiceanu CM, Caffo BS, et al. **Population-wide principal component-based quantification of blood-brain-barrier dynamics in multiple sclerosis.** *NeuroImage* 2011;57:1430–46