



**Providing Choice & Value**

Generic CT and MRI Contrast Agents



**FRESENIUS  
KABI**

**CONTACT REP**

**AJNR**

**Systematic Assessment of Multispectral  
Voxel-Based Morphometry in Previously  
MRI-Negative Focal Epilepsy**

R. Kotikalapudi, P. Martin, J. Marquetand, T. Lindig, B.  
Bender and N.K. Focke

This information is current as  
of July 28, 2025.

*AJNR Am J Neuroradiol* 2018, 39 (11) 2014-2021

doi: <https://doi.org/10.3174/ajnr.A5809>

<http://www.ajnr.org/content/39/11/2014>

# Systematic Assessment of Multispectral Voxel-Based Morphometry in Previously MRI-Negative Focal Epilepsy

R. Kotikalapudi, P. Martin, J. Marquetand, T. Lindig, B. Bender, and N.K. Focke



## ABSTRACT

**BACKGROUND AND PURPOSE:** Voxel-based morphometry is widely used for detecting gray matter abnormalities in epilepsy. However, its performance with changing parameters, smoothing and statistical threshold, is debatable. More important, the potential yield of combining multiple MR imaging contrasts (multispectral voxel-based morphometry) is still unclear. Our aim was to objectify smoothing and statistical cutoffs and systematically compare the performance of multispectral voxel-based morphometry with existing T1 voxel-based morphometry in patients with focal epilepsy and previously negative MRI.

**MATERIALS AND METHODS:** 3D T1-, T2-, and T2-weighted FLAIR scans were acquired for 62 healthy volunteers and 13 patients with MR imaging negative for focal epilepsy on a Magnetom Skyra 3T scanner with an isotropic resolution of 0.9 mm<sup>3</sup>. We systematically optimized the main voxel-based morphometry parameters, smoothing level and statistical cutoff, with T1 voxel-based morphometry as a reference. As a next step, the performance of multispectral voxel-based morphometry models, T1+T2, T1+FLAIR, and T1+T2+FLAIR, was compared with that of T1 voxel-based morphometry using gray matter concentration and gray matter volume analysis.

**RESULTS:** We found the best performance of T1 at 12 mm and a T-threshold (statistical cutoff) of 3.7 for gray matter concentration analysis. When we incorporated these parameters, after expert visual interpretation of concordant and discordant findings, we identified T1+FLAIR as the best model with a concordant rate of 46.2% and a concordant rate/discordant rate of 1.20 compared with T1 with 30.8% and 0.67, respectively. Visual interpretation of voxel-based morphometry findings decreased concordant rates from 38.5%–46.2% to 15.4%–46.2% and discordant rates from 53.8%–84.6% to 30.8%–46.2% and increased specificity across models from 33.9%–40.3% to 46.8%–54.8%.

**CONCLUSIONS:** Multispectral voxel-based morphometry, especially T1+FLAIR, can yield superior results over single-channel T1 in focal epilepsy patients with a negative conventional MR imaging.

**ABBREVIATIONS:** AUC = area under the curve; C<sub>R</sub> = concordant rate; D<sub>R</sub> = discordant rate; EEG = electroencephalography; GMC = gray matter concentration; GMV = gray matter volume; S<sub>p</sub> = specificity; VBM = voxel-based morphometry

In focal epilepsy, detection of a focal lesion in MR imaging increases the odds of seizure-free outcome after an operation by 2.5–3 times.<sup>1</sup> A common epileptogenic lesion is focal cortical dys-

plasia, which has notable morphologic characteristics on MR imaging.<sup>2</sup> However, pure visual analysis, especially in subtle cases, can be challenging. Also, many histopathologically proven focal cortical dysplasias escape visual detection.<sup>3</sup> Failure to identify these lesions can often label patients as having negative MR imaging for focal cortical dysplasia. These patients have poor surgical outcomes or may not even be referred for a potential epilepsy surgery.<sup>4</sup>

During the past decade, multiple MR imaging postprocessing methods have been applied to improve lesion detection in epilepsy,<sup>5</sup> of which the most common applications are based on voxel-based morphology (VBM), usually using T1-weighted images (single-channel T1 VBM). VBM enables a voxelwise comparison between 2 groups of subjects and can highlight areas of statistically significant differences.<sup>6</sup> This approach is commonly applied in the presurgical evaluation of epilepsy, comparing a single patient against a group of healthy controls. Martin et al, in 2015,<sup>5</sup> reviewed various studies that have reported the potential use of VBM based on T1

Received February 5, 2018; accepted after revision August 6.

From the Departments of Diagnostic and Interventional Neuroradiology (R.K., T.L., B.B.) and Neurology and Epileptology (R.K., P.M., J.M., N.K.F.), Hertie Institute for Clinical Brain Research, University Hospital Tübingen, Tübingen, Germany; and Department of Clinical Neurophysiology (R.K., N.K.F.), University Hospital Göttingen, Göttingen, Germany.

This work was funded by the University of Tübingen/Applied Clinical Research (321-0-1) and Deutsche Forschungsgemeinschaft (FO750/5-1).

Please address correspondence to Raviteja Kotikalapudi, Department of Clinical Neurophysiology, University Hospital Göttingen, Robert-Koch-Str 40, 37075 Göttingen, Germany; e-mail: raviteja.kotikalapudi@gmail.com

Indicates open access to non-subscribers at [www.ajnr.org](http://www.ajnr.org)

Indicates article with supplemental on-line tables.

Indicates article with supplemental on-line photos.

<http://dx.doi.org/10.3174/ajnr.A5809>

within a sensitivity range of 60%–100%, whereas its application in identifying structural abnormalities in cryptogenic epilepsy remains challenging, with a sensitivity between 10% and 38%. Also, past studies used variable smoothing (between 5 and 14 mm)<sup>7–9</sup> and statistical cutoffs to elucidate VBM findings.<sup>10,11</sup> Studies have shown that changing smoothing kernel size<sup>9</sup> and statistical thresholds have a direct impact on the VBM findings,<sup>10</sup> also affecting the specificity of the results.<sup>10</sup>

Recently, multispectral VBM (ie, the combination of multiple MR imaging contrast sequences) was proposed and made available in SPM8/SPM12 (<http://www.fil.ion.ucl.ac.uk/spm/software/spm12>) within its unified segmentation framework (New Segmentation). This approach enables combining different MR image contrasts. Lindig et al, in 2018,<sup>12</sup> showed that multispectral VBM improves tissue segmentation for gray matter, white matter, and CSF and improves lesion detection in MR imaging positive for focal epilepsy compared with T1 VBM. However, it is still unclear whether multispectral VBM is beneficial in the more challenging cohort of MR imaging negative for epileptogenic lesions.

In this study, we wanted to objectify the selection of smoothing and statistical cutoffs with reference to the established and frequently used T1 VBM. This enabled us to systematically compare the diagnostic value of T1 and multispectral VBM using multiple image contrast combinations, T1+T2, T1+FLAIR, and T1+T2+FLAIR, in patients with focal epilepsy, but negative conventional MR imaging (ie, previously MRI-negative).

## MATERIALS AND METHODS

### Subjects

We recruited 62 healthy volunteers (number of controls,  $N_c$ ; 36 women, 26 men; mean age, 27.5 years) and 13 patients with MR imaging negative for cryptogenic epilepsy (number of patients,  $N_p$ ; 6 women, 7 men; mean age, 35.9 years) with a lobar clinical hypothesis of epilepsy origin. The clinical hypothesis was established through multiexpert consensus in the monthly epilepsy case conference, which uses all available clinical information in the presurgical epilepsy program (On-line Table 1). The multispectral VBM results analyzed in this work were not used in forming the clinical hypothesis. All patients had undergone video-electroencephalography (EEG) telemetry and neuropsychological assessment, and 6 of them had undergone PET/CT. After an epilepsy case conference for each of the 13 patients, 7 were identified as candidates for further intracranial EEG and possible subsequent epilepsy surgery due to medically refractory seizures (for clinical details see On-line Table 1). To date, intracranial EEG was performed in 5 patients and was offered to 2 patients who have not yet decided. Of the 5 patients with implants, 3 patients (patients 1, 2, and 6) had undergone an operation; in 2 patients with implants (patients 4 and 11), surgery was declined due to suspected multifocal epileptogenesis. In postsurgical patients, histopathology revealed evidence of hippocampal sclerosis in 1 patient (patient 2), but this patient did not achieve postoperative seizure freedom (Engel class II). The histopathologic tests of the other 2 operated patients (patients 1 and 6) were unremarkable, and to date, 1 patient (patient 1) is seizure-free without anticonvulsive medication (Engel I). The other patient (patient 6) had postoperative seizure freedom, but only for 1 year. All subjects were scanned on a 3T MR

imaging scanner (Magnetom Skyra; Siemens, Erlangen, Germany) at the University Hospital Tübingen. The acquisition protocol was previously reported<sup>12</sup> and consists of 0.9-mm<sup>3</sup> isotropic 3D T1-weighted MPRAGE (= T1), 3D T2-weighted sampling perfection with application-optimized contrasts by using different flip angle evolutions (SPACE sequence; Siemens) (= T2), and 3D inversion recovery prepared SPACE sequence (= FLAIR).

### Image Processing

At first, all scans were converted from DICOM to NIfTI format using MRIConvert (<https://lcn.uoregon.edu/downloads/mriconvert>). The postprocessing was performed with SPM12 in a Matlab R2016a (MathWorks, Natick, Massachusetts) environment and with an in-house Matlab code. We used VBM based on T1 images only (T1 VBM) as a reference. This was compared with multispectral VBM combinations based on multiple MR images (ie, T1+T2, T1+FLAIR, and T1+T2+FLAIR). For multispectral VBM, T2 images (T2 and FLAIR channels) were coregistered to their respective T1 images using linear coregistration with 12 *df* and a normalized mutual information cost function. Segmentation was performed with default settings of bias regularization of 0.001 and a bias cutoff full width at half maximum of 60 mm. As a next step, the segmented GM images were normalized to the Montreal Neurological Institute space by using the Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL, part of SPM) toolbox, with an isotropic resolution of 1.5 mm<sup>3</sup>.<sup>13</sup> To correct for regional volume changes in normalization, we modulated images to preserve the quantity of GM tissue (gray matter volume [GMV]) within a voxel and unmodulated images for preserving the concentration of GM tissue (gray matter concentration [GMC]). Finally, spatially normalized images were smoothed, using a Gaussian kernel with a range from 4- to 16-mm full width at half maximum in a step size of 2. The smoothed GMC and GMV images were further analyzed using SPM12.

### Statistical Analysis

We used the general linear model from SPM12 to analyze regional increased differences in the smoothed GMC and GMV images. The general linear model analysis was performed using the factorial design specification (2-sample *t* test) in SPM12 to compare each patient against the control cohort (patient comparison) and each control in a leave-one-out cross-validation against the remainder of the controls after removing the subject in question (control comparison). We included age and sex as covariates for GMC. Total intracranial volume was also included as a covariate for GMV. The resulting statistical maps (*t*-contrast maps, subject > controls, ie, increase of GM) were thresholded with *T*-score cutoffs from 2.5 to 6 in step sizes of 0.1. The suprathreshold clusters at each smoothing level and each *T*-threshold were used for reporting the results.

### Objective Diagnostic Performance Assessment

We created a brain mask by summation of normalized unmodulated GM and WM images from T1 segmentation. Voxel intensities <0.5 were excluded to remove nonbrain areas. On the basis of the expert clinical hypothesis, the lobes of the hypothesis were identified for each patient in the Montreal Neurological Institute structural atlas provided with FSL, Version 5.0

(<http://www.fmrib.ox.ac.uk/fsl>).<sup>14,15</sup> For each patient, “concordant lobes” were defined as the atlas lobes/regions (for example, left frontal lobe) that are identified as the lobes of clinical hypothesis. The lobes that are not a part of the clinical hypothesis for each patient were defined as “discordant lobes.” For controls, because no epileptogenic lesions are expected, all atlas regions were defined as discordant.

### Analysis

Every suprathreshold cluster was considered as a concordant or a discordant finding provided one-third or greater of the suprathreshold cluster overlapped the respective lobar ROI (concordant or discordant).

The concordant rate ( $C_R$ ) was calculated as

$$C_R = \left( \frac{N_{CON}}{N_P} \right) \times 100.$$

The discordant rate ( $D_R$ ) was calculated as

$$D_R = \left( \frac{N_{DIS}}{N_P} \right) \times 100.$$

For controls, each suprathreshold cluster was considered a non-epileptogenic finding provided one-third or greater of the cluster overlapped the control cortical mask (excluding likely artifactual findings outside the brain). Specificity ( $S_p$ ) was calculated as

$$S_p = \left( \frac{N_{NF}}{N_C} \right) \times 100.$$

$N_{CON}$ ,  $N_{DIS}$  and  $N_{NF}$  in the above equations refer to the number of patients with concordant and discordant findings and the number of controls with no findings, respectively.

### Diagnostic Assessment

To determine the ideal parameters (smoothing level and T-threshold), we generated receiver operating characteristic curves at each smoothing level by plotting  $100 - S_p$  versus  $C_R$  for all T-thresholds. The area under the curve (AUC) was calculated using a trapezoidal integration function in Matlab as a performance index for each smoothing kernel width. For comparing single-versus-multispectral VBM, we considered the smoothing level with the best performance (the highest AUC across all smoothing levels; 4–16 mm) for T1 as a reference.

### Determining T-Threshold, $S_p$ , and $C_R$

For this reference smoothing kernel,  $S_p$  and  $C_R$  values at T-thresholds of 2.5–6 were plotted.<sup>8</sup> We considered the T-threshold in which the remainder of  $C_R - S_p$  was the smallest but still positive. This point was referred to as the “optimized T-threshold.” At this threshold, a balanced trade-off between these 2 diagnostic indices ( $C_R$  and  $S_p$ ) can be achieved. At this defined T-threshold and the reference smoothing level,  $S_p$ ,  $C_R$ , and  $D_R$  values were reported for all VBM models. Later, the Euclidean distance (ED) (Fig 1D) of the pair  $S_p$ ,  $C_R$  from (100, 100) was calculated for all VBM models as

$$ED = \sqrt{(100 - S_p)^2 + (100 - C_R)^2}.$$

### Visual Interpretation of VBM Findings

On the basis of the minimum Euclidean distance value across models and analysis, we selected analysis findings for visual interpretation. Visual analysis was performed to verify VBM findings in patients and controls. At the reference smoothing kernel and the optimized T-threshold, all findings across models were inverse-transformed to native space using the deformation utility in SPM12 for each subject separately (“back-normalization”). To group clusters in a close spatial relation, we applied a smoothing of 8-mm full width at half maximum to the transformed cluster maps and considered all clusters connected at a statistical value of  $>0.5$  as a single cluster for the visual review process. Later, these clusters (native space findings) across all models were combined for each subject. Finally, native space findings were overlaid on corresponding patient native T1 and coregistered T2 and FLAIR scans. An expert board-certified neuroradiologist (B.B.) inspected each finding without prior knowledge of the lobar hypothesis and labeled each cluster as the following: 1, potentially epileptogenic and visible; 2, potentially epileptogenic but not visible; 3, nonepileptogenic; 4, unclear/not visible or ambiguous; and 5, artifacts. For visual inspection of controls, each cluster was categorized into the following: 1, visible and nonepileptogenic; 2, unclear/not visible; and 3, artifacts. The reviewer used the non-epileptogenic label when the finding was visible but likely not epileptogenic, such as microangiopathy or perivascular spaces. VBM clusters were rated as unclear when the finding was not sufficiently visible to confirm these findings as potentially epileptogenic, nonepileptogenic, or artifacts. As part of the visual analysis, we reviewed all clusters again to ascertain the correctness of the automated lobar classification and to flag them as either concordant or discordant clusters. This step was needed in only 1 case in which 63.2% of voxels were in the concordant lobar mask. Hence, this cluster was eventually marked as a concordant finding.

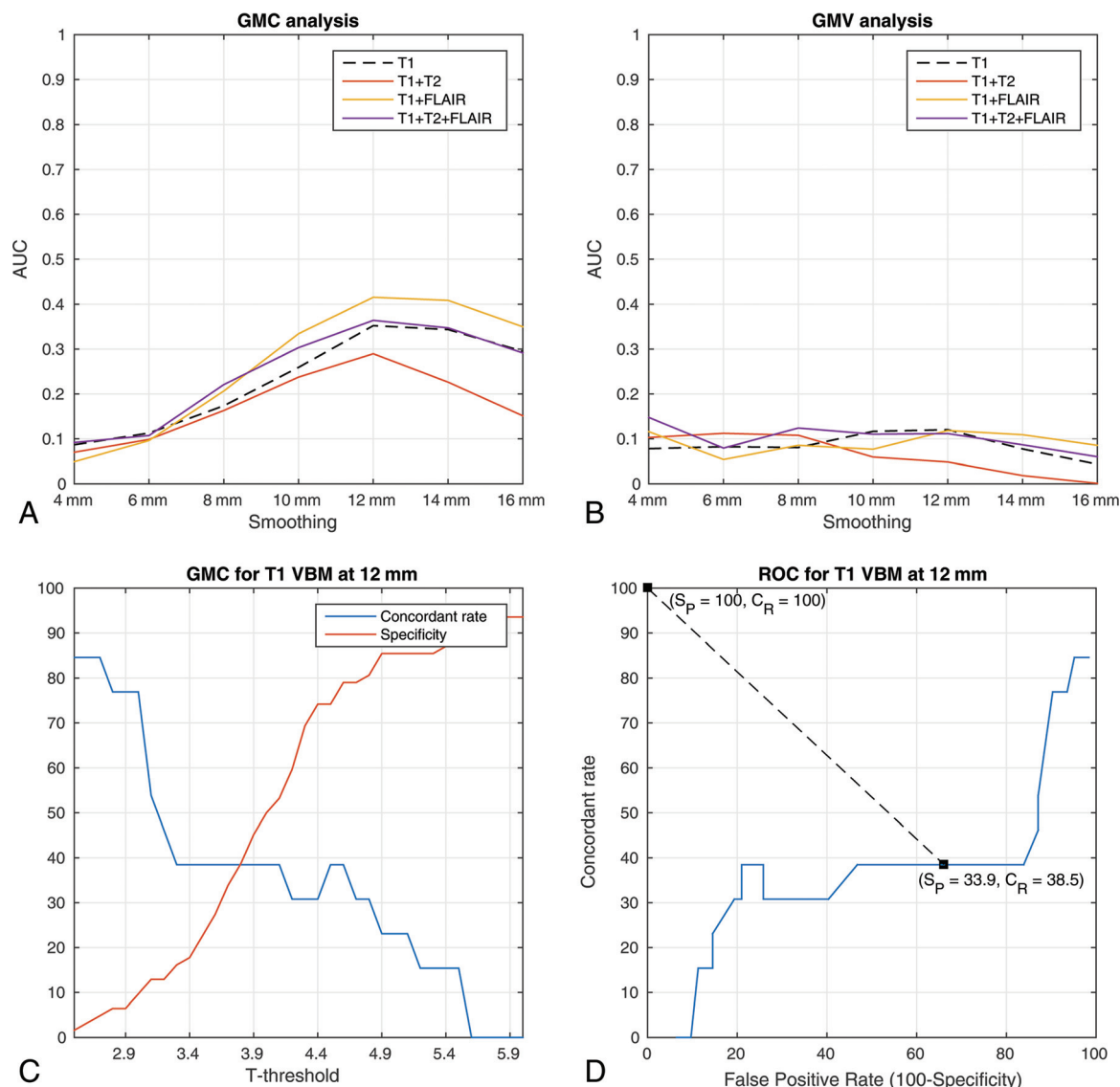
## RESULTS

### VBM Automated Results: Smoothing Parameters in GMC and GMV

For all models in the GMC analysis, the ideal smoothing was found to be 12 mm. T1+FLAIR showed the best AUC of 0.42 at 12 mm in comparison with T1 (0.35), T1+T2 (0.29), and T1+T2+FLAIR (0.36). For the GMV analysis, the ideal smoothing was also at 12-mm full width at half maximum of different smoothing levels as shown in Fig 1A, -B and On-line Table 2.

### T-Threshold, Concordant Rate, and Specificity

We found the intersection of  $C_R$  and  $S_p$  across different T-thresholds at 3.7 for 12 mm in GMC analysis (Fig 1C). At this T-threshold, T1+T2 and T1+FLAIR showed  $C_R$ ,  $S_p$  at 46.2, 40.3 and 46.2, 37.1, respectively, compared with T1 at 38.5, 33.9. The ratio of  $C_R/D_R$  was also higher for T1+T2 and T1+FLAIR at 0.86 and 0.75, respectively, compared with T1. The Euclidean distances for all models were in the range of 80.37–90.29. The performance details of all models are presented in Table 1. For the GMV analysis, a lower T-threshold of 3.0 was found at 12 mm as the intersection point for  $C_R$  and  $S_p$ . The concordant rate across models was between 7.7% and 38.5%, with specificity between 9.7% and



**FIG 1.** Diagnostic performance of different VBM models. A and B, The area under the curve for different smoothing kernels for gray matter concentration and volume. C, An example of an intersection plot for T1 VBM for the concordant rate and specificity against statistical cutoffs. D, An example of a receiver operating characteristic curve for T1 VBM at 12-mm smoothing for GMC. The Euclidean distance is calculated from the optimized T-threshold (3.7 in GMC analysis), where  $C_R$ ,  $S_P = 38.5, 33.9$  to  $C_R$ ,  $S_P = 100, 100$ . False positive rate =  $100 - S_P$ .

**Table 1: GMC analysis for VBM models at 12 mm and a T-threshold of 3.7<sup>a</sup>**

Model	Concordant Rate (%)	Specificity (%)	Discordant Rate (%)	Euclidean Distance (Range, 0–141.14)	Concordant Rate/Discordant Rate
T1	38.5	33.9	84.6	90.29	0.46
T1+T2	46.2	40.3	53.8	80.37	0.86
T1+FLAIR	46.2	37.1	61.5	82.77	0.75
T1+T2+FLAIR	46.2	35.5	76.9	84.00	0.60

<sup>a</sup>For VBM GMC automated analysis (prior to visual interpretation), values of concordant rate, specificity, discordant rate, Euclidean distance from  $C_R = S_P = 100$ , and concordant/discordant ratio are provided for all models: namely, T1, T1+T2, T1+FLAIR, and T1+T2+FLAIR. The smoothing level and T-threshold are 12 mm and 3.7, respectively.

21%.  $C_R/D_R$  across models was between 0.10 and 0.50. The Euclidean distances were in the range of 105.32–123.66. All details for GMV analysis are provided in Table 2.

### Visual Interpretation Results of VBM Findings

Visual interpretation was performed for GMC analysis at 12-mm full width at half maximum and a T-threshold of 3.7. The highest concordant rate was found for T1+FLAIR at 46.2% in compari-

son with T1 at 30.8%. Also, T1+FLAIR showed the highest  $C_R/D_R$  ratio at 1.20 compared with T1 (0.67) and the rest of the models. Only T1+FLAIR and T1+T2+FLAIR showed more or equal concordant-to-discordant findings. The discordant finding rate dropped from 53.8%–84.6% to 30.8%–46.2% after visual interpretation. All results for visual interpretation of patient findings are presented in Table 3. In the visual interpretation of controls, 14.5%–21% of findings were marked as nonepileptogenic, while



**Table 2: GMV analysis for VBM models at 12 mm and a T-threshold of 3.0<sup>a</sup>**

Model	Concordant Rate (%)	Specificity (%)	Discordant Rate (%)	Euclidean Distance (Range, 0–141.14)	Concordant Rate/Discordant Rate
T1	23.1	21.0	84.6	110.25	0.27
T1+T2	7.7	17.7	76.9	123.66	0.10
T1+FLAIR	38.5	9.7	84.6	109.25	0.46
T1+T2+FLAIR	38.5	14.5	76.9	105.32	0.50

<sup>a</sup> For VBM GMV automated analysis (prior to visual interpretation), values of concordant rate, specificity, discordant rate, Euclidean distance from  $C_R = S_P = 100$ , and concordant/discordant ratio are provided for all models: namely, T1, T1+T2, T1+FLAIR, and T1+T2+FLAIR. The smoothing level and T-threshold are 12 mm and 3.0, respectively.

**Table 3: Patients—visual analysis results for GMC<sup>a</sup>**

Model	Potentially Epileptogenic (and Visible)	Potentially Epileptogenic (and Not Visible)	Potentially Epileptogenic (Combined)	Nonepileptogenic	Unclear	Artifacts
Concordant lobe (%)						
T1	15.4	15.4	30.8	7.7	15.4	0
T1+T2	15.4	0	15.4	7.7	15.4	0
T1+FLAIR	15.4	30.8	46.2	7.7	23.1	0
T1+T2+FLAIR	15.4	15.4	30.8	7.7	23.1	0
Discordant lobe (%)						
T1	7.7	38.5	46.2	7.7	30.8	15.4
T1+T2	7.7	23.1	30.8	0	23.1	0
T1+FLAIR	7.7	30.8	38.5	7.7	30.8	7.7
T1+T2+FLAIR	7.7	23.1	30.8	0	30.8	7.7

<sup>a</sup> Visual interpretation results of GMC analysis of patients are provided. The results contain percentages of patients scored by the reviewer as potentially epileptogenic and visible, potentially epileptogenic and not visible, potentially epileptogenic (number of patients with potentially epileptogenic and visible/not visible or both), nonepileptogenic, unclear, and artifacts. The results are reported for both concordant and discordant lobes.

**Table 4: Controls—visual analysis results for GMC<sup>a</sup>**

Models	All Lobes (%)			Corrected Specificity (Excluding Artifacts and Nonepileptogenic)
	Nonepileptogenic	Unclear	Artifacts	
T1	14.5	53.2	22.6	46.8
T1+T2	21.0	45.2	17.7	54.8
T1+FLAIR	16.1	50.0	19.4	50.0
T1+T2+FLAIR	16.1	48.4	25.8	51.6

<sup>a</sup> Visual interpretation results for GMC analysis for controls are presented. Results contain the percentage of controls scored by the reviewer as nonepileptogenic, unclear, and artifacts. Finally, corrected specificity is reported as the percentage of controls that did not have unclear findings—that is, all findings identified as artifacts/nonepileptogenic lesions and patients with no findings (VBM specificity prior to visual analysis).

45.2%–53.2% were marked as unclear (Table 4); 17.7%–25.8% of controls also had findings that were classified as artifacts. The overall specificity after visual interpretation across models was between 46.8% and 54.8%.

## DISCUSSION

In this study, we identified ideal analysis parameters, namely smoothing and T-threshold (statistical cutoff), with reference to T1 VBM. Incorporating these parameters, we systematically compared T1 and multispectral VBM using a combination of T1, T2, and FLAIR images for detecting gray matter structural abnormalities in patients with MR imaging previously negative for focal epilepsy. We objectified the main VBM parameters, smoothing and statistical cutoffs, with reference to the classic T1 VBM. Furthermore, by systematically comparing multispectral VBM models against T1 VBM and using visual inspection of VBM findings, we found T1+FLAIR as the best performing model based on the concordant rate, specificity, and concordant-to-discordant ratio.

## Smoothing and Statistical Cutoff

There is a wide variation in parameters in VBM studies for lesion detection in epilepsy for smoothing kernel sizes ranging from 5 to 14 mm<sup>7-9,16-18</sup> and for a statistical cutoff from  $P < .001$  uncorrected (corresponding to a statistical cutoff T of approximately 3.2 in our sample) to  $P < .05$  family wise error–corrected (corresponding to a statistical cutoff T of approximately 5.1 in our sample)<sup>10</sup> as well as comparisons at multiple statistical cutoffs.<sup>8</sup> We provide a systematic comparison of the 2 main parameters focused on patients with

focal epilepsy, but negative conventional MRI, given that this is the main target cohort for advanced lesion detection in epilepsy. For the most frequently used T1-only approach, we found that a smoothing of 12-mm full width at half maximum and a threshold of  $T = 3.7$  provide the highest AUC and best diagnostic balance between concordant rate and specificity. Most interesting, our recommendation of a 12-mm smoothing kernel is similar to that in a previous study conducted by Salmond et al.<sup>19</sup> In line with this study, we found a low specificity with decreased diagnostic performance (AUC, Fig 1A and On-line Table 2) at 4 mm across all models.

Second, as expected, the concordant rate decreased and specificity increased with respect to increasing statistical cutoffs (Fig 1C). This finding is in agreement with a previous study on a lesional cohort based on Z-scores.<sup>8</sup> We aimed for a balance between sensitivity/ $C_R$  and specificity (maximal AUC, intersection of  $C_R$  and  $S_P$ ). It is debatable whether a different approach with maximized sensitivity could also be clinically useful, dealing with little or no specificity. However, in our view, a VBM approach should have at least some

degree of specificity to be informative in the difficult context of MRI-negative patients, in which invasive diagnostics and invasive EEG are commonly performed and carry low-but-non-negligible risks for patients. Martin et al<sup>10</sup> found that the best odds ratio for predicting postoperative seizure freedom was achieved by the VBM variant with the best specificity, namely the normalized FLAIR signal. Nonetheless, our results will also provide guidance if other groups want to select their analysis parameters with a different intention, either maximizing sensitivity or specificity (On-line Figs 1 and 2).

### Gray Matter Concentration versus Volume Analysis

We found that gray matter concentration analysis revealed better results for both single- and multispectral models in comparison with volume analysis. This is in agreement with a previous study in a different cohort in which gray matter concentration was found to be better in detecting epileptogenic lesions in MRI-negative and MRI-positive patients.<sup>10</sup> The only difference between the 2 analyses is the additional step of modulation, which is intended to preserve the original tissue volume.<sup>6</sup> The exact reason for this diagnostic difference remains speculative. It is possible that epileptogenic lesions and focal cortical dysplasias are better characterized by local changes of tissue composition, and global volume effects attenuate the effect size (eg, “compensating” gray matter increase in the lesions with atrophy in the surrounding area). Whether superiority of GMC over GMV analysis remains true for different MR imaging pulse sequences with different tissue contrasts needs to be determined.

### Visual Interpretation of VBM Findings

After visual interpretation of control findings, we found a specificity range between 46.8% and 54.8% across all models. We cannot draw a direct comparison with previous studies because our VBM models, smoothing and statistical cutoffs, are different from those in past studies.<sup>10,16,20</sup> However, we are in line with our previous study on 50 controls, in which a specificity of 42% was reported for T1 gray matter concentration analysis<sup>10</sup> at  $P < .05$  (family-wise error). Moreover, 12%–16.1% of controls had only visible (nonepileptogenic/artifacts) findings (ie, controls who had visible findings but no unclear findings). This finding contributed to an increase in specificity after visual analysis (Table 4). In a previous study based on the morphometric analysis program, approximately 25% of findings in controls (13 of 52) were marked as nonepileptogenic normal variants,<sup>20</sup> similar to our results. Across models, after visual analysis, we observed a drop in concordant and discordant rates (Tables 1 and 3). In a previous study based on an MR imaging cohort negative for epileptogenic lesions, after visual analysis, the concordant rate for FLAIR and T1 VBM dropped from 28.6% and 14.3% to 14.3% and 0%, respectively.<sup>21</sup> This result shows that visual interpretation of VBM findings is needed before considering these for epilepsy surgery. Nevertheless, we and others have previously shown that VBM findings carry a positive odds ratio for a good outcome after epilepsy surgery and are, thus, clinically relevant.<sup>10,20</sup>

### Effect of Multispectral VBM

After visual inspection, the highest concordant rate was confirmed for multispectral T1 + FLAIR at 46.2% in comparison with T1 (30.8%) and the other models. This is in line with a previous study in patients with lesional epilepsy, in which multispectral

**Table 5: Nonvisible findings—visual analysis results for GMC<sup>a</sup>**

Models	All Lobes (%)	
	Controls	Patients
T1	53.2	61.5
T1 + T2	45.2	30.8
T1 + FLAIR	50.0	61.5
T1 + T2 + FLAIR	48.4	46.2

<sup>a</sup> Rates of findings not visible—that is, unclear findings in controls and unclear/potentially epileptogenic and not visible findings in patients across all models for all lobes (%) are reported.

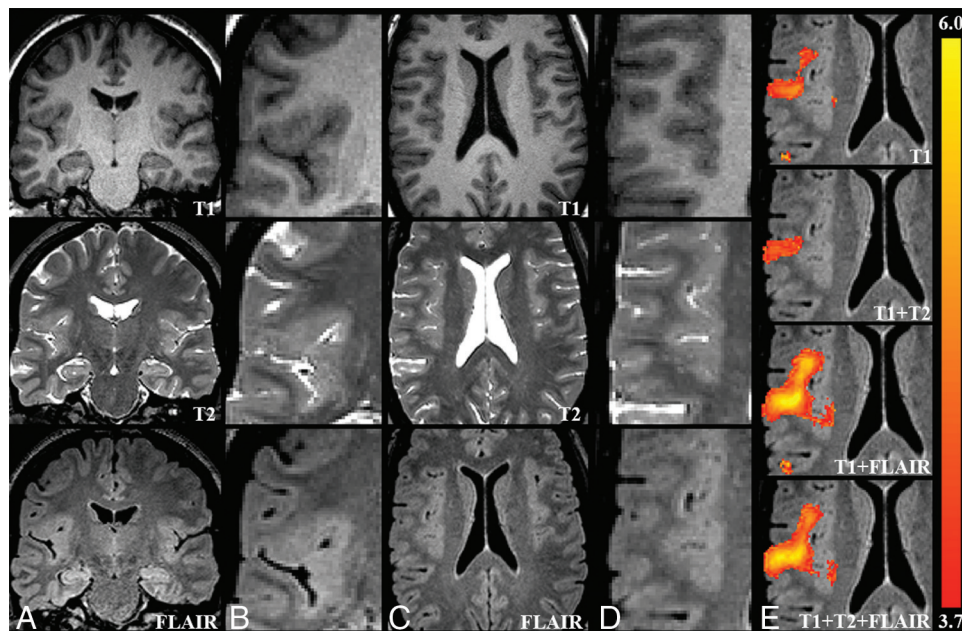
T1 + FLAIR also showed a superior performance compared with T1.<sup>12</sup> Similarly, the enhanced performance of FLAIR-VBM (concordant rate, 14.3%) over T1 (0%) was also found in a study cohort of children with cryptogenic epilepsy.<sup>21</sup> Additionally, intensity-normalized FLAIR-VBM<sup>11</sup> showed positive odds of 7.3 for a successful surgical outcome in comparison with T1 VBM in 129 patients with MR imaging negative for epilepsy.<sup>10</sup> This finding indicates that inclusion of a new channel, FLAIR in this case, can contribute to increasing concordant rates in a cohort with MR imaging negative for focal epilepsy. Overall these results show that multispectral VBM is superior to T1 VBM, but there still is a need for new modalities and approaches to be explored (eg, based on [resting-state] functional imaging, alternative tissue contrasts like MP2RAGE/diffusion imaging, and higher magnetic fields).

### Discordant Findings in Patients

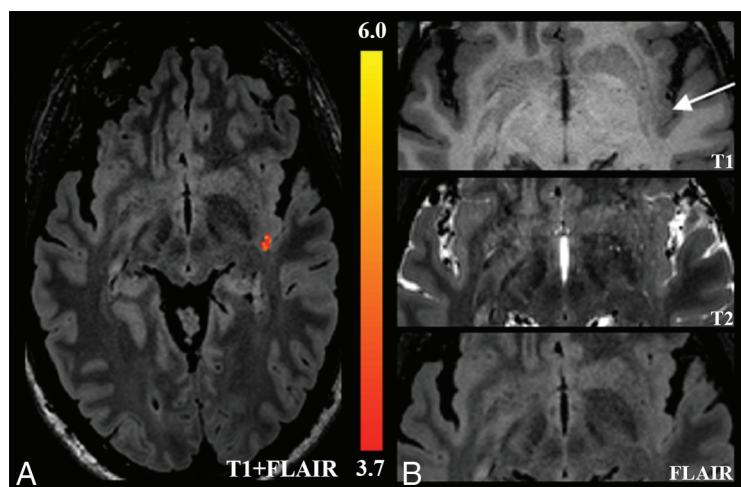
All VBM variants had a relevant number of findings discordant with the primary clinical hypothesis, but these differed substantially among the variants. Only T1 + FLAIR showed more concordant-than-discordant findings, with the highest concordant-to-discordant ratio among all models and approximately 2 times more than in T1. In a previous study, Martin et al<sup>10</sup> showed that only normalized FLAIR-VBM had more concordant than discordant findings in MRI-negative patients against T1. In the same study, only normalized FLAIR-VBM had no discordant findings in the group of 15 patients with MR imaging positive for focal cortical dysplasia. We are not aware of any other VBM-based study addressing discordant findings in the cryptogenic epilepsy cohort; 30.8%–46.2% of patients still had discordant findings that were marked as potentially epileptogenic. In this most challenging epilepsy surgery cohort (MR imaging negative for focal epilepsy), the definition of an epileptogenic zone is often limited because fast propagation of epileptogenic activity can be difficult to detect in scalp video-EEG.<sup>22,23</sup> Furthermore, patients can have >1 epileptogenic zone/lesion, and our hypotheses were derived from noninvasive data (especially video telemetry EEG) in many cases (On-line Table 1). Thus, VBM findings initially considered discordant may still be real and clinically relevant. We also provided the concordant and discordant findings in patients for all VBM models that were visually confirmed as potentially epileptogenic as in On-line Table 3.

### Nonvisible Findings

Patients and controls also had findings marked as not visible (Table 5). The biologic meaning of these findings that are not visible, is difficult to assess. They can be due to subtle artifacts beyond the visible threshold (eg, field inhomogeneities or movement) or minor differences in cortical morphology without direct pathologic mean-



**FIG 2.** Sample case (patient 4). *A* and *C*, An overview in the native space T1, T2, and FLAIR images. *B* and *D*, Zoomed-in images focusing on a possible epileptogenic lesion. A subtle disruption of cortical morphology is visible in this figure, more prominent on FLAIR images. This finding is concordant with the clinical hypothesis supported by noninvasive and intracranial EEG, indicating seizure onset in the right frontal lobe. The possible lesion was detected as increased GM by all VBM models, but the extent and effect size were clearly better for the VBM models, including FLAIR (*E*).



**FIG 3.** Sample case (patient 13). *A*, A VBM finding detected only by T1+FLAIR in the left temporoinsular region overlaid on the native space FLAIR image. *B*, Magnified images of T1, T2, and FLAIR suggesting a blurred gray-white matter junction (arrow). This finding is concordant with the clinical hypothesis of seizure onset.

ing. In patients, these findings can also represent subtle epileptogenic lesions that escape visual detection. In a direct comparison, the frequency of detecting a nonvisible finding was only slightly higher in patients than in controls (for some models), making it likely that these findings are largely nonspecific. However, more clinical data and follow-up are needed to draw further conclusions from such nonvisible findings.

### Diagnostic Value of Multispectral FLAIR-VBM in Focal Epilepsy

A major advantage of multispectral VBM lies in simultaneously including FLAIR along with T1. It takes only an extra 5–7 minutes

to acquire 3D-FLAIR. Additionally, processing time for T1 and T1+FLAIR VBM differs only in the extra step of coregistration for the latter. Second, many lesions in focal epilepsy (focal cortical dysplasia, hippocampal sclerosis, tumors) have a prolonged T2-relaxation, resulting in hyperintense signal and increased visibility on a FLAIR image.<sup>2,21</sup> Thus, it is not surprising that FLAIR-based methods could detect those lesions better than T1 alone.<sup>11,12</sup> For example, Fig 2 shows the presumed lesion to be more visible on FLAIR than on T1 images. All VBM models eventually detect this abnormality, but the coverage and effect size were best in VBM models that include FLAIR. The finding is in concordance with the clinical lobar hypothesis, which was in the right frontal lobe, indicated by noninvasive EEG and intracranial EEG.

Another example is shown in Fig 3, where only T1+FLAIR VBM detects a subtle structural abnormality in the left temporoinsular region. In comparison with the right side, a subtle blurring of GM-WM is visible at the left temporal operculum/posterior insula. This finding is in concordance with the clinical lobar hypothesis of bilateral temporal onset. In both cases, it is likely that the suspected abnormalities are segmented as gray matter and, consequently, increase the gray matter probability when the FLAIR image is added. Furthermore, the improved tissue classification by multimodal segmentation may improve the lesion detection.<sup>12,24,25</sup> Of the 3 patients with operations, 1 patient (patient 6) had findings concordant with the lobar clinical hypothesis in the left temporal region detected by T1



and T1+FLAIR but not coinciding with the resected area. The patient was seizure-free for only 1 year before seizures resumed, raising the possibility of a partial resection and prompting further clinical investigations.

### Limitations

Five of 6 patients with concordant findings did not have surgical resection; the 1 patient with a resection did not have the finding in the resection area and later had seizure relapse. Hence, a histopathologic confirmation was not possible. Low rates of surgical resection as well as reduced chances of seizure freedom are typical for patients with MR imaging negative for epilepsy.<sup>1</sup> Nonetheless, a visual review of the findings by an expert radiologist blinded to the clinical hypothesis shows clear improvement in the concordant/discordant ratio for multispectral VBM. This indicates that these findings can be of diagnostic relevance. However, our study cannot assess whether multispectral VBM will eventually improve the surgical outcome. This will require longitudinal multimodal studies by large multicenter collaborations. Our results can guide the choice of sequences and analysis parameters for such projects.

### CONCLUSIONS

We provide optimized VBM parameters, specifically smoothing and statistical cutoff (T-threshold), for lesion detection in MR imaging previously negative for focal epilepsy with T1 VBM as a reference. We systematically compared multispectral VBM with T1 VBM. We found a smoothing level of 12 mm and a T-threshold of 3.7 for GMC analysis as ideal parameters. With these parameters, after we visually confirmed VBM findings, multispectral VBM T1+FLAIR yielded results superior to those of all other models. We also found VBM an important computational advancement, which, after a careful visual interpretation, can aid the presurgical evaluation of focal epilepsy. We recommend multispectral VBM, especially T1+FLAIR, as currently the best VBM model in detecting increased gray matter structural abnormalities in patients with MR imaging previously negative for focal epilepsy.

Disclosures: Benjamin Bender—UNRELATED: Expert Testimony: Medtronic\*. Niels K. Focke—RELATED: Grant: Deutsche Forschungsgemeinschaft/Deutsche Forschungsgemeinschaft, Comments: Deutsche Forschungsgemeinschaft grant FO750/5–1; UNRELATED: Consultancy: Bial, Eisai; Payment for Lectures Including Service on Speakers Bureaus: UCB; Payment for Development of Educational Presentations: Bial. \*Money paid to the institution.

### REFERENCES

1. Téllez-Zenteno JF, Hernández Ronquillo L, Moien-Afshari F, et al. **Surgical outcomes in lesional and non-lesional epilepsy: a systematic review and meta-analysis.** *Epilepsy Res* 2010;89:310–18 CrossRef Medline
2. Blümcke I, Thom M, Aronica E, et al. **The clinicopathologic spectrum of focal cortical dysplasias: a consensus classification proposed by an ad hoc Task Force of the ILAE Diagnostic Methods Commission.** *Epilepsia* 2011;52:158–74 CrossRef Medline
3. Wang ZI, Alexopoulos AV, Jones SE, et al. **The pathology of magnetic-resonance-imaging-negative epilepsy.** *Mod Pathol* 2013;26:1051–58 CrossRef Medline
4. Spencer SS, Berg AT, Vickrey BG, et al; Multicenter Study of Epilepsy Surgery. **Predicting long-term seizure outcome after resective epilepsy surgery: the multicenter study.** *Neurology* 2005;65:912–18 CrossRef Medline
5. Martin P, Bender B, Focke NK. **Post-processing of structural MRI for individualized diagnostics.** *Quant Imaging Med Surg* 2015;5:188–203 CrossRef Medline
6. Ashburner J, Friston KJ. **Voxel-based morphometry: the methods.** *Neuroimage* 2000;11(6 Pt 1):805–21 CrossRef Medline
7. Huppertz HJ, Grimm C, Fauser S, et al. **Enhanced visualization of blurred gray-white matter junctions in focal cortical dysplasia by voxel-based 3D MRI analysis.** *Epilepsy Res* 2005;67:35–50 CrossRef Medline
8. Colliot O, Bernasconi N, Khalili N, et al. **Individual voxel-based analysis of gray matter in focal cortical dysplasia.** *Neuroimage* 2006;29:162–71 CrossRef Medline
9. Keller SS, Roberts N. **Voxel-based morphometry of temporal lobe epilepsy: an introduction and review of the literature.** *Epilepsia* 2008;49:741–57 CrossRef Medline
10. Martin P, Winston GP, Bartlett P, et al. **Voxel-based magnetic resonance image postprocessing in epilepsy.** *Epilepsia* 2017;58:1653–64 CrossRef Medline
11. Focke NK, Symms MR, Burdett JL, et al. **Voxel-based analysis of whole brain FLAIR at 3T detects focal cortical dysplasia.** *Epilepsia* 2008;49:786–93 CrossRef Medline
12. Lindig T, Kotikalapudi R, Schweikardt D, et al. **Evaluation of multimodal segmentation based on 3D T1-, T2- and FLAIR-weighted images: the difficulty of choosing.** *Neuroimage* 2018;170:210–21 CrossRef Medline
13. Ashburner J. **A fast diffeomorphic image registration algorithm.** *Neuroimage* 2007;38:95–113 CrossRef Medline
14. Mazziotta J, Toga A, Evans A, et al. **A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM).** *Philos Trans R Soc Lond B Biol Sci* 2001;356:1293–322 CrossRef Medline
15. Collins DL, Holmes CJ, Peters TM, et al. **Automatic 3-D model-based neuroanatomical segmentation.** *Hum Brain Mapp* 1995;3:190–208 CrossRef
16. Focke NK, Bonelli SB, Yogarajah M, et al. **Automated normalized FLAIR imaging in MRI-negative patients with refractory focal epilepsy.** *Epilepsia* 2009;50:1484–90 CrossRef Medline
17. Kassubek J, Huppertz HJ, Spreer J, et al. **Detection and localization of focal cortical dysplasia by voxel-based 3-D MRI analysis.** *Epilepsia* 2002;43:596–602 CrossRef Medline
18. Bonilha L, Montenegro MA, Rorden C, et al. **Voxel-based morphometry reveals excess gray matter concentration in patients with focal cortical dysplasia.** *Epilepsia* 2006;47:908–15 CrossRef Medline
19. Salmond C, Ashburner J, Vargha-Khadem F, et al. **Distributional assumptions in voxel-based morphometry.** *Neuroimage* 2002;17:1027–30 CrossRef Medline
20. Wang ZI, Jones SE, Jaisani Z, et al. **Voxel-based morphometric magnetic resonance imaging (MRI) postprocessing in MRI-negative epilepsies.** *Ann Neurol* 2015;77:1060–75 CrossRef Medline
21. Riney CJ, Chong WK, Clark CA, et al. **Voxel based morphometry of FLAIR MRI in children with intractable focal epilepsy: implications for surgical intervention.** *Eur J Radiol* 2012;81:1299–305 CrossRef Medline
22. Alarcón G, Kissani N, Dad M, et al. **Lateralizing and localizing values of ictal onset recorded on the scalp: evidence from simultaneous recordings with intracranial foramen ovale electrodes.** *Epilepsia* 2001;42:1426–37 Medline
23. Spencer SS, Williamson PD, Bridgers SL, et al. **Reliability and accuracy of localization by scalp ictal EEG.** *Neurology* 1985;35:1567–75 CrossRef Medline
24. Viviani R, Pracht ED, Brenner D, et al. **Multimodal MEMPRAGE, FLAIR, and R2\* segmentation to resolve dura and vessels from cortical gray matter.** *Front Neurosci* 2017;11:258 CrossRef Medline
25. Viviani R, Stöcker T, Stingl JC. **Multimodal FLAIR/MPRAGE segmentation of cerebral cortex and cortical myelin.** *Neuroimage* 2017;152:130–41 CrossRef Medline