AJNR

Voxel-Based Morphometry—from Hype to Hope. A Study on Hippocampal Atrophy in Mesial Temporal Lobe Epilepsy

F. Riederer, R. Seiger, R. Lanzenberger, E. Pataraia, G. Kasprian, L. Michels, J. Beiersdorf, S. Kollias, T. Czech, J. Hainfellner and C. Baumgartner

Voxel-Based Morphometry—from Hype to Hope. A Study on Hippocampal Atrophy in Mesial Temporal Lobe Epilepsy

F. Riederer, R. Seiger, R. Lanzenberger, E. Pataraia, G. Kasprian, L. Michels, J. Beiersdorf, S. Kollias, T. Czech, J. Hainfellner, and C. Baumgartner

ABSTRACT

BACKGROUND AND PURPOSE: Automated volumetry of the hippocampus is considered useful to assist the diagnosis of hippocampal sclerosis in temporal lobe epilepsy. However, voxel-based morphometry is rarely used for individual subjects because of high rates of false-positives. We investigated whether an approach with high dimensional warping to the template and nonparametric statistics would be useful to detect hippocampal atrophy in patients with hippocampal sclerosis.

MATERIALS AND METHODS: We performed single-subject voxel-based morphometry with nonparametric statistics within the framework of Statistical Parametric Mapping to compare MRI from 26 well-characterized patients with temporal lobe epilepsy individually against a group of 110 healthy controls. The following statistical threshold was used: $P < .05$ corrected for multiple comparisons with family-wise error over the region of interest right and left hippocampus.

RESULTS: The sensitivity for the detection of atrophy related to hippocampal sclerosis was 0.92 (95% CI, 0.67–0.99) for the right hippocampus and 0.60 (0.31–0.83) for the left, and the specificity for volume changes was 0.98 (0.93–0.99). All clusters of decreased hippocampal volumes were correctly lateralized to the seizure focus. Hippocampal volume decrease was in accordance with neuronal cell loss on histology reports.

CONCLUSIONS: Nonparametric voxel-based morphometry is sensitive and specific for hippocampal atrophy in patients with mesial temporal lobe epilepsy and may be useful in clinical practice.

ABBREVIATIONS: TLE = temporal lobe epilepsy; VBM = voxel-based morphometry; TIV = total intracranial volume; ILAE = International League Against Epilepsy; DARTEL = Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra; MAP = Morphometric Analysis Program; ROC = Receiver operating characteristic curves; AUC = area under the curve; SnPM = Statistical non Parametric Mapping; HS = hippocampal sclerosis; FLAIR = fluid attenuated inversion recovery; EEG = electroencephalography

Voxel-based morphometry (VBM) is a powerful automated tool to investigate cerebral gray matter changes, based on high-resolution structural MR imaging.[1] It has been used to demonstrate network atrophy in temporal lobe epilepsy (TLE)[2-4] and neuroplasticity in pain conditions, such as medication-overuse headache,[5] and in learning.[6]

While inference from VBM studies is mostly based on group comparisons, information for the individual subject would be of great clinical value. The use of VBM for single subjects however is limited by a high rate of false-positive findings.[7] Recent studies demonstrate that this problem can be overcome with the use of nonparametric statistics.[8]

TLE is the most frequent among the focal epilepsies. The most common underlying pathology in patients with temporal lobe epilepsy is mesial temporal sclerosis characterized by selective hippocampal neuronal loss and gliosis.[9] In medically refractory TLE, hippocampal atrophy is associated with a favorable

outcome after epilepsy surgery[10] and in a large epilepsy surgery series, hippocampal sclerosis was the most frequent finding.[9] Thus, identification of hippocampal sclerosis (HC) is of high clinical relevance during presurgical evaluation. A previous study found that VBM was not sensitive to cortical neuronal loss and hippocampal sclerosis in individual patients.[11] Voxel-based automated analyses have rarely been used for the detection of hippocampal sclerosis in the individual.[12-14] Recent methodologic advances, such as the use of a high-dimensional warping to the template brain, are likely to improve sensitivity of VBM.[15]

The aim of our study was to evaluate the potential role of VBM for the detection of hippocampal atrophy in individual epilepsy patients, by using high-dimensional warping and nonparametric statistics in a clinical setting. This was done by comparing well-characterized individual patients with known hippocampal sclerosis against a large normal data base from healthy controls, estimating sensitivity and specificity. Specifically, we studied whether subjects with known hippocampal atrophy and healthy controls were correctly assigned by the automated algorithm.

MATERIALS AND METHODS

Overview

The study was approved by the ethics committee of the Medical University of Vienna. The conduct of the study is in accordance with the Declaration of Helsinki. In brief, hippocampal volumes were estimated automatically, based on high-resolution 3D structural MR imaging scans. Hippocampal volumes were assessed by 1) comparing individual MR imaging scans against a large normal data base from healthy controls, by using VBM and nonparametric statistics that automatically detect voxels with gray matter changes; 2) plotting estimates of extracted hippocampal volumes from individual patients with temporal lobe epilepsy against the healthy control data base.

Patients with Epilepsy and Healthy Control MR Imaging Data Base

The present retrospective analysis includes high-resolution MR imaging scans from 26 patients with TLE (5 patients from the Karl Landsteiner Institute of Clinical Epilepsy Research, acquired on a 3T Achieva scanner (Philips), termed "3T VIE," and 21 patients from the Medical University Vienna, acquired on a 1.5T Gyroscan (Philps), termed "1.5T VIE" from a previous study that are fully described therein),[2] fulfilling the following criteria (all inclusion criteria and none of the exclusion criteria): 1) Patients with drug-resistant epilepsy, ie, those who had a failure of adequate trials of 2 tolerated, appropriately chosen and used antiepileptic drug schedules (monotherapy or combination) to achieve persistent seizure freedom.[16] 2) Diagnosis of TLE defined by typical clinical seizure semiology, interictal and ictal EEG findings documented during video-EEG monitoring. 3) Patients had undergone high-resolution T1-weighted volume MR imaging sequences of the brain (eg, MPRAGE). Exclusion criteria: 1) Lesions in the brain, except hippocampal sclerosis. 2) Significant vascular comorbidity or vascular risk factors such as uncontrolled hypertension. Demographic and clinical data including MR imaging findings are summarized in the On-line Table. Mean age

was 38.2 ± 13.0 years (range 18–60), 13 were women. Definition of the criterion standard (hippocampal sclerosis, side) was based on the following: An expert panel consisting of neurologists, neuroradiologists, and neurosurgeons reached a consensus on the diagnosis and the side of seizure focus, based on clinical semiology, ictal and interictal electroencephalography (EEG), and a dedicated MR imaging epilepsy protocol. Histology data and postsurgical outcome were available in the 20/26 patients who had had an operation.

Healthy controls were recruited by using local advertisements at the research sites and screened by senior neurologists. Our healthy control data base comprised high-resolution 3D T1-weighted sequences from the Karl Landsteiner Institute of Clinical Epilepsy Research and Cognitive Neurology, Vienna (acquired on "3T VIE," n = 44) and from previous studies at the Institute for Biomedical Engineering, University and ETH Zurich (acquired on a 3T Achieva scanner, termed "3T ZH," n = 66).[17,18] In addition, high-resolution MR imaging scans acquired on "1.5T VIE" from the previous epilepsy study[2] were available, n = 12. Thus, the healthy control data base comprised a total of 122 subjects (71 women, 41 men; mean age 36.2 ± 12.5 years, range 22–62). The size of the control group was based on previous work on nonparametric single-subject VBM.[7,8]

Details of the scanning protocols are given in the On-line Appendix. All images were examined by a senior neuroradiologist and only participants without structural abnormalities or motion artifacts were included. The inclusion of data from different scanners was considered helpful for the evaluation of the method in a real-world setting, where scanners may occasionally be replaced.

Demographic, Clinical, and Follow-Up Data

Demographic and clinical data and radiologic MR imaging reports were acquired from electronic charts. Follow-up data from the patients from the previous epilepsy study[2] were collected if available, by using electronic charts from the Medical University of Vienna, including histology (if the patient had undergone epilepsy surgery) and seizure outcome according to the International League Against Epilepsy (ILAE) classification,[19] considering seizure control during the last available 12 months. Mean follow-up duration of operated patients was 9.6 ± 4.1 years. Detailed histologic reports were collected from operated patients and retrospectively assigned according to the ILAE classification scheme for hippocampal sclerosis (HS)[20] based on description of cell loss in the respective subfields.

VBM and Estimation of Hippocampal Gray Matter Volumes

Images were segmented into gray matter, white matter, and cerebrospinal fluid with the VBM8 toolbox (http://dbm.neuro.uni-jena.de/wordpress/vbm/download/), incorporated in the SPM8 (http://www.fil.ion.ucl.ac.uk/spm/software/spm8/) software running on Matlab R2008b (MathWorks). The VBM8 toolbox uses a high-dimensional Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL) for optimized registration of different brains[15] to the template brain as a default setting. The basic idea behind DARTEL is to increase the accuracy
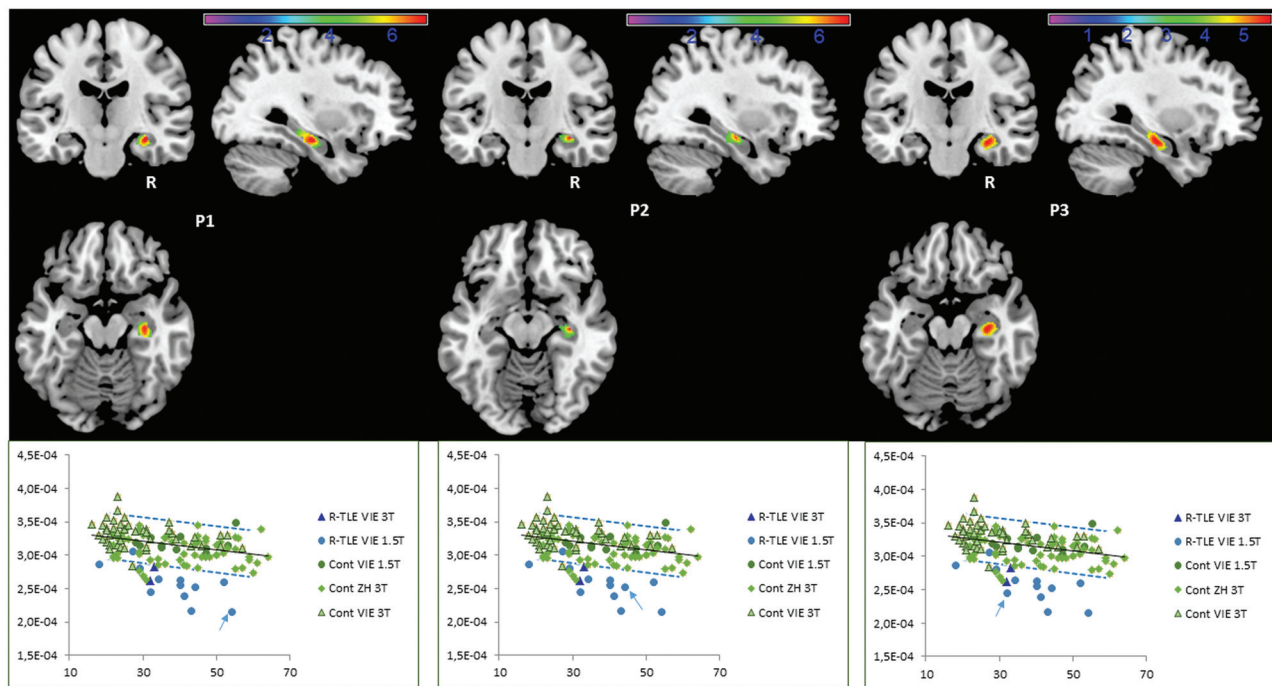
FIG 1. Comparison of representative individual patients with right-sided TLE against a normal data base of 110 healthy controls. The color scale indicates the nonparametric *t*-value. Results are displayed at $P < .05$ corrected for multiple comparisons with family-wise error across the ROI. The scatterplot shows extracted hippocampal volumes/total intracranial volume for controls and patients (y-axis, institutional units), where the individual patient is indicated with a blue arrow. Different symbols indicate the control cohorts. The x-axis indicates the age. Blue dotted lines show percentiles 5 and 95, respectively.

of intersubject alignment by modeling the shape of each scan by using millions of parameters (3 per voxel).[15] The default DARTEL-template in VBM8 was used to keep the analysis pipeline identical for all subjects, independent from the scanner site. The sampling distance was set to 1 and the bias regularization to extremely light, expecting high-intensity nonuniformity artifact; otherwise default settings were used. The "modulation" procedure ensures the preservation of volume information after normalization to the template. The normalized GM segments were smoothed by using a 6-mm full width at half maximum Gaussian kernel. Individual MR imaging scans of patients were compared against the healthy controls from "3T ZH" and "3T VIE" ($n = 110$), by using Statistical non Parametric Mapping[21] (SnPM, http://www2.warwick.ac.uk/fac/sci/statistics/staff/academic-research/nichols/software/snpm), as this approach is less dependent on normality of data and is less likely to yield false-positives. SnPM is a toolbox for SPM based on nonparametric permutation testing[21] where normality of data and equal variance are not required. This should be an advantage in unbalanced statistical designs such as single-case VBM where one subject is compared against a control group.[8] Age, sex, and total intracranial volume (TIV), which is calculated as the sum of GM, white matter, and cerebrospinal fluid by the VBM toolbox, were used as nuisance variables.

Only voxels in the right or left hippocampus (as defined by the MarsBaR toolbox http://marsbar.sourceforge.net/about.html) were considered. The following threshold was defined a priori: $P < .05$, corrected for multiple comparisons with family-wise error, across the ROI right + left hippocampus (small volume correction).

The proportion of patients with known hippocampal atrophy who had clusters of decreased GM in the respective hippocampus was calculated (sensitivity). In a second step, in an identical way, individual MRI of the healthy control group ($n = 122$) was successively compared against the healthy controls from "3T ZH" and "3T VIE" ($n = 110$), to determine specificity (true-negative rate = 1 – false-positive rate). Thus, sensitivity and specificity of the volumetric approach for the detection of hippocampal atrophy was calculated with 95% Wilson confidence intervals.

Subsequently, GM of the right and left hippocampus was extracted automatically from individual normalized and smoothed GM segments by using the MarsBaR toolbox (http://marsbar.sourceforge.net/). Individual hippocampal volumes for the right and left sides divided by TIV were plotted against the data from healthy controls, from which percentiles 5 and 95 were estimated. The hippocampal volumes in subjects with known hippocampal sclerosis would be expected to be below percentile 5 of controls, based on pilot data.

Finally, for quality control, test-retest reliability for extracted hippocampal volumes was calculated based on scans of healthy controls who were scanned twice at "3T ZH." Possible interscanner differences were estimated, comparing right and left hippocampal volumes of healthy controls between the 3 scanners by using MarsBaR.

RESULTS
VBM in Individual Patients and Specificity of Findings
Significant volume decreases are shown in Fig 1 and On-line Fig 1. Results for individual patients including nonparametric *t*-

Sensitivity and specificity of single-subject VBM for the detection of hippocampal atrophy

Sensitivity for the detection of hippocampal atrophy in patients with mesial temporal lobe epilepsy[a]				
	R Hippocampus decrease	L Hippocampus decrease		
	0.92 (0.67–0.99)	0.60 (0.31–0.83)		
Specificity of hippocampal volume changes				
	R Hippocampus		L Hippocampus	
	Decrease	Increase	Decrease	Increase
Specificity	0.99 (0.96–1.00)	1.00 (0.97–1.00)	0.99 (0.96–1.00)	0.99 (0.96–1.00)
	Decrease or increase in R or L Hippocampus			
Specificity	0.98 (0.93–0.99)			

[a] Wilson 95% CI is given in brackets.

scores are given in the On-line Table. The sensitivity for the detection of hippocampal sclerosis was 0.92 for the right side and 0.60 for the left, with a high specificity of 0.98 (Table). Importantly, right-sided hippocampal atrophy could be detected in 1 patient (P13) in whom routine MR imaging was considered normal (only fluid attenuated inversion recovery [FLAIR] hyperintensity was detected with the automated Morphometric Analysis Program [MAP]).[22] In 4 patients with left-sided hippocampal sclerosis, no atrophy could be detected automatically at this threshold, and 2 of them had only very mild signs of hippocampal sclerosis on MR imaging (On-line Table, clinical data). All clusters of decreased hippocampal volumes were on the side of the seizure focus.

Extracted Hippocampal Volumes in Patients with Temporal Lobe Epilepsy

In 12/13 patients with hippocampal sclerosis/atrophy on the right side, extracted hippocampal volumes/TIV were below percentile 5 ipsilaterally (Fig 1), corresponding to a sensitivity of 0.92 (95% Wilson CI, 0.67–0.99). The patient with a normal MR imaging report (P13) and FLAIR hyperintensity in automated postprocessing on the right side had a hippocampal volume/TIV just below percentile 5. Conversely, in 8/10 patients with hippocampal sclerosis/atrophy on the left side, hippocampal volumes/TIV were below percentile 5, corresponding to a sensitivity of 0.80 (0.49–0.94). The 3 cases with left hippocampal volumes/TIV above percentile 5 had no or only very discrete signs of atrophy in radiologic reports. Lateralization was correct in all cases (with low hippocampal volume), ie, the side of low hippocampal volume/TIV always corresponded to the side of seizure onset. Receiver operating characteristic curves (ROC) for the detection of hippocampal atrophy based on extracted hippocampal volumes for the right and left hippocampus are shown in On-line Fig 2. The area under the curve (AUC) was 0.964 for the right hippocampus and 0.900 for the left hippocampus, indicating excellent accuracy. ROC analyses including only subjects from the scanner 1.5T VIE showed similar AUC values (right hippocampus 0.993; left hippocampus 0.898).

Histology from Resected Hippocampal Specimens and Postsurgical Outcome

Histopathologic results and postsurgical outcome according to the ILAE classification[19] were available in all patients who underwent epilepsy surgery (n = 20; On-line Table). Radiologic findings of hippocampal sclerosis were confirmed in all operated

cases histologically (On-line Table) as a finding of partial HS (type 1 in most cases). One case with FLAIR hyperintensity in the hippocampus and signs of volume increase in automated volumetry (P25) had no hippocampal sclerosis, but did have discrete dispersion of granular cells and reactive oligodendral hyperplasia. Seventeen out of 19 cases with hippocampal sclerosis on histology showed decreased hippocampal volume on automated volumetry. Outcome was favorable in most cases (On-line Table), in accordance with the literature.[10]

Hippocampal Volumes in Healthy Controls

Left hippocampal volumes were significantly larger compared with right hippocampal volumes (0.47 ± 0.046 versus 0.44 ± 0.046; $P < .001$, paired t-test). Left and right hippocampal volumes/TIV (ratio) showed a significant negative correlation with age (left: $r = -0.196$, $P = .031$; right: $r = -0.361$, $P < .001$).

Test-Retest Reliability

Thirty-seven healthy controls were scanned twice at the 3T ZH, with a mean time interval of 166.4 ± 27.9 days. Extracted hippocampal volumes/TIV were stable between time points, as evidenced by good correlations for right ($r = 0.95$) and left hippocampal volumes ($r = 0.96$; On-line Fig 3). In addition, hippocampal volumes were compared between time points on a voxel basis, by using a paired t-test design within SPM. At the whole brain level, corrected for multiple comparisons with family-wise error, no significant differences were found. In ROI analyses, a few voxels showed decreased gray matter in the region of the right hippocampal tail at the later time point (5 voxels at x = 26, y = −36, z = 7 and 1 voxel at x = 23, y = −36, z = 4; On-line Fig 3).

Comparisons between Scanners

Hippocampal volumes were compared between healthy controls scanned with the scanners 3T VIE, 3T ZH, and 1.5T VIE, by using an ANOVA model within SPM and the MarsBar toolbox, with the covariates age, sex, and TIV. The Bonferroni correction for multiple testing was applied. For the left hippocampus, no significant differences were found between the scanners. Right hippocampal volumes from the 3T ZH were significantly lower compared with those from the 3T VIE ($P < .001$, corrected) and 1.5T VIE ($P = .001$, corrected). In age- and sex-matched healthy controls, right hippocampal volume was 7.1% lower on data from 3T ZH as compared with 3T VIE. However, these between scanner effects for the right hippocampus were considerably lower than the differences related to hippocampal atrophy in patients (On-line Fig 4).

DISCUSSION

Various VBM techniques have been evaluated for the detection of occult brain lesions in focal epilepsy.[23] The present study focuses on the detection of hippocampal atrophy based on automated analyses of T1-weighted MR imaging of the brain. Our study demonstrates that hippocampal atrophy related to mesial temporal sclerosis can be automatically detected in individual patients by VBM by using nonparametric statistics with high sensitivity and specificity. Using small volume correction for multiple comparisons, the finding of atrophy was highly specific (0.99), indicating that this procedure adequately minimizes false-positives. Furthermore, VBM always correctly lateralized the epileptogenic zone. Our data indicate that this technique should be useful in clinical practice, because it enables an objective user-independent measurement of individual atrophy and is less time-consuming than manual volumetric methods. Likewise, various automated hippocampal volumetry techniques including the use of FreeSurfer (http://surfer.nmr.mgh.harvard.edu) and NeuroQuant (CorTech Labs, San Diego, California) have been proposed to increase the diagnostic sensitivity for hippocampal sclerosis.[24-30] In one using NeuroQuant, hippocampal asymmetry z scores had the best sensitivity (86.7%–89.5%) and specificity (92.2%–94.1%) to discriminate patients with TLE from healthy controls with AUC values ranging from 0.915–0.939.[27] In another study, the combination of automated hippocampal T2-relaxometry techniques with volumetry has been shown to improve separation of patients with HS from healthy controls.[31] Similarly, an automated volumetry and FLAIR analysis tool improved the diagnosis of ambiguous HS.[32] Huppertz et al[33] showed that an automated quantitative FLAIR analysis in SPM could assist the diagnosis of hippocampal sclerosis with high sensitivity (97.1%) and specificity (95.4%).

In contrast to these studies that primarily analyzed estimates of hippocampal volumes, the present work is focused on automated voxel-based analyses, providing topographic information on atrophy patterns. Voxel-based techniques have rarely been used for the individual diagnosis of hippocampal sclerosis. Bonilha et al[12] demonstrated voxelwise standardized z scores to be helpful for the detection of hippocampal sclerosis (with an AUC of 0.973) and more recently, machine-learning techniques have been used to discriminate right or left mesial TLE from healthy controls.[13,14] Our findings contrast with one previous study that found VBM not to be sensitive for the detection of hippocampal sclerosis and neuronal loss in individual patients.[11] Methodologic issues such as improved preprocessing by using high-dimensional warping, the use of nonparametric statistics, and a large control group probably can explain this discrepancy. One previous study on healthy controls showed that the high rate of false-positives frequently observed in single-subject VBM studies[7] can be minimized by using nonparametric statistics, which is not based on the assumption of normal distribution and equal variance.[8] By comparing 122 individual healthy controls successively against the healthy control group, we confirmed a low false-positive rate (2%), ie, high specificity (0.98), as long as correction for multiple testing was applied. Conversely, correction for multiple comparisons across the whole brain may be too conservative, when the main interest lies on hippocampal structures and we propose an a priori-defined small volume correction to be

optimal in this case. Extracting hippocampal volumes and comparing these to normal values can further strengthen the detection of atrophy in individual patients.

As right and left hippocampal volumes significantly differed in favor of the left side in healthy controls, it is mandatory to estimate normal ranges separately for both sides. Asymmetry of mesial temporal lobe structures (left > right), including the hippocampal formation, has been found in an earlier VBM study on a large group of healthy controls.[34] Although, this may depend on analysis techniques, because the reverse asymmetry has also been reported by using NeuroQuant[27] or FreeSurfer.[24] Right and left hippocampus showed only a slight decrease with age in contrast to an obvious total gray matter volume loss (data not shown), corroborating previous findings indicating that the hippocampus may be relatively spared from volume loss over a wide age range in healthy controls.[35]

For quality control, additional analyses in healthy controls were conducted, considering reproducibility of volumetric measurements and possible variability between MR imaging scanners. As 37 healthy controls were scanned twice with the 3T ZH, test-retest reliability could be calculated and was considered excellent, with r-values ranging from 0.95–0.96 for extracted hippocampal volumes. Voxel-based comparisons between time points showed no significant differences between time points at whole brain level; however, ROI analyses for the right and left hippocampus revealed a few voxels with reduced gray matter in the region of the right hippocampal tail at the later time point. These were considered unlikely to have an influence on the results because of the distinct anatomic localization.

The present study used MR imaging data from 3 different scanners, two 3T scanners of the same type (3T VIE and 3T ZH, both Achieva) and one 1.5T scanner from the same manufacturer (1.5T VIE, Gyroscan). Motivation for this approach was the assumption that a large control group would ultimately increase the sensitivity for alterations in individual patients. In addition, a large control group of $n = 198$ was also used in the study of Scarpazza et al[8] on nonparametric VBM. Furthermore, MR imaging scanners are being replaced from time to time, so that data from different machines may be collected when a rare condition is to be studied over a longer time period. Finally, multicenter MR imaging studies may be an emerging strategy in the study of rare disorders. In our study, a scanner effect was seen only for the right hippocampus with decreased volume estimates for data from 3T ZH. This possibly decreased the sensitivity for the detection of right HS, because a large portion of the healthy controls was investigated at this site. Ultimately, interscanner effects could not be exactly determined, because different subjects were scanned in Vienna and in Zurich. However, in the between-scanners analyses, sex and age were used as nuisance variables, so that these finding should not be confounded by demographic variability. Age- and sex-matched comparisons of mean hippocampal volumes confirmed lower hippocampal volumes in subjects scanned with 3T ZH compared with 3T VIE (7.1% lower); whereas comparisons with 1.5T VIE were not significant after correction for multiple comparisons, but this sample was rather small ($n = 11$, data not shown).

The variability of automated volumetric measurements between scanners has previously been estimated to range up to 14.7%

(coefficient of variation) depending on the structures studied.[36] Our data indicate, however, that the volumetric changes detected in hippocampal sclerosis clearly outweigh interscanner differences (Online Fig 4). Thus, combining data from different scanners may be feasible when the expected disease-related effects are expected to exceed the interscanner variability. Interestingly, intersite effects for various MR imaging measures including gray matter volume have been observed even after harmonizing imaging protocols.[37] Nevertheless, harmonization of scanning protocols and equipment should be attempted whenever possible.

In the present study, hippocampal sclerosis could be verified by histology in most cases. This suggests in vivo determined volume decrease to be associated with cell loss in the case of hippocampal sclerosis, though this could not be shown directly in our study. Histologic sections were not available for quantitative analysis and thus, no correlations between MR imaging derived hippocampal volume and histologic measures such as cell size or attenuation could be performed. Hippocampal subfield volumes estimated by a manual segmentation approach have been shown to correlate with neural attenuation and size derived from histology.[38] Likewise, automatically estimated total hippocampal volumes were correlated with histopathologic neural attenuation in the hippocampus and its subfields.[39]

Most of the operated cases were considered to have type 1 HS, which is the most abundant form with good prognosis.[20,40] Our data indicate that automated detection of volume decrease related to HS is probably not confined to this subtype as a few subjects with type 2 or 3 HS also showed significant atrophy.

Previous studies have consistently shown atrophy in patients with mesial TLE beyond the hippocampus in regions anatomically and functionally related, such as thalamus and cingulum, ie, network atrophy,[2-4,41] based on group comparisons. The nonparametric approach of the present study may have the potential to identify network atrophy on an individual level though sensitivity and specificity of findings would have to be thoroughly assessed across regions. Thus, the prevalence of atrophy in particular regions could be estimated, which is probably of interest for our understanding of TLE.[42] This was, however, beyond the scope of our study, and further work on this issue will be necessary.

As a limitation of the present study, it has to be mentioned that the sensitivity and specificity of findings apply exclusively to the hippocampus that has been analyzed; findings may differ for other regions in terms of false-positives,[7] depending also on region-specific smoothness of data. The present study did not evaluate the proposed analysis pipeline for discrimination between mesial TLE and other forms of epilepsy. Another limitation is the pooling of MR imaging data from different sites without harmonization of study protocols and equipment, inherent in the retrospective design. Further, in the present study, no direct comparisons between hippocampal volume and histologic measures could be performed. Finally, it has to be acknowledged that hippocampal volume loss is only one of the hallmarks of hippocampal sclerosis, where hyperintensity on FLAIR and T2-weighted images, loss of internal structure, as well as changes on diffusion tensor imaging may also be considered.[25,43] It should also be emphasized that the presented analysis pipeline is meant to assist neuroradiologists¨ and clinicians¨ work in the context of dedicated radiologic and clinical evaluation.

Strengths of the present study include the well-characterized group of patients with epilepsy with histologic and long-term follow-up data and the large number of healthy controls being subsequently and individually compared with the control group. SPM and its toolboxes are freely available, which could encourage the clinical application of VBM for volumetry. Required computational resources were considered within acceptable limits because processing time per image was approximately 50 minutes.

CONCLUSIONS

Although reproducibility and specificity of findings from VBM studies have been a matter of debate, our findings show that nonparametric VBM is sensitive and specific for hippocampal atrophy in patients with mesial TLE, and we encourage its use in clinical practice. Further, we provide evidence that MR imaging data from different centers may be combined for morphometric studies if gross structural abnormalities are expected. The importance of harmonization of MR imaging scanning protocols and equipment should be considered in future prospective studies. Lastly, because VBM can provide information on whole-brain topographic changes, future studies could investigate individual atrophy patterns in TLE.

ACKNOWLEDGMENTS

The authors express their gratitude to the developers of Statistical Parametric Mapping at the Wellcome Trust Center for Neuroimaging, United Kingdom, and Statistical non Parametric Mapping, at the University of Warwick, United Kingdom, for distributing their software packages to the community.

REFERENCES

1. Ashburner J, Friston KJ. **Voxel-based morphometry–the methods.** *NeuroImage* 2000;11:805–21 CrossRef Medline
2. Riederer F, Lanzenberger R, Kaya M, et al. **Network atrophy in temporal lobe epilepsy: a voxel-based morphometry study.** *Neurology* 2008;71:419–25 CrossRef Medline
3. Bonilha L, Rorden C, Castellano G, et al. **Voxel-based morphometry reveals gray matter network atrophy in refractory medial temporal lobe epilepsy.** *Arch Neurol* 2004;61:1379–84 CrossRef Medline
4. Mueller SG, Laxer KD, Cashdollar N, et al. **Voxel-based optimized morphometry (VBM) of gray and white matter in temporal lobe epilepsy (TLE) with and without mesial temporal sclerosis.** *Epilepsia* 2006;47:900–07 CrossRef Medline
5. Riederer F, Gantenbein AR, Marti M, et al. **Decrease of gray matter volume in the midbrain is associated with treatment response in medication-overuse headache: possible influence of orbitofrontal cortex.** *J Neurosci* 2013;33:15343–49 CrossRef Medline

6. Draganski B, Gaser C, Busch V, et al. **Neuroplasticity: changes in grey matter induced by training.** *Nature* 2004;427:311–12 CrossRef Medline

7. Scarpazza C, Sartori G, De Simone MS, et al. **When the single matters more than the group: very high false positive rates in single case voxel based morphometry.** *NeuroImage* 2013;70:175–88 CrossRef

8. Scarpazza C, Nichols TE, Seramondi D, et al. **When the single matters more than the group (ii): addressing the problem of high false positive rates in single case voxel based morphometry using non-parametric statistics.** *Front Neurosci* 2016;10:6 CrossRef Medline

9. Blumcke I, Spreafico R, Haaker G, et al; EEBB Consortium. **Histopathological findings in brain tissue obtained during epilepsy surgery.** *N Engl J Med* 2017;377:1648–56 CrossRef Medline

10. Jones AL, Cascino GD. **Evidence on use of neuroimaging for surgical treatment of temporal lobe epilepsy: a systematic review.** *JAMA Neurol* 2016;73:464–70 CrossRef Medline

11. Eriksson SH, Thom M, Symms MR, et al. **Cortical neuronal loss and hippocampal sclerosis are not detected by voxel-based morphometry in individual epilepsy surgery patients.** *Hum Brain Mapp* 2009;30:3351–60 CrossRef Medline

12. Bonilha L, Halford JJ, Rorden C, et al. **Automated MRI analysis for identification of hippocampal atrophy in temporal lobe epilepsy.** *Epilepsia* 2009;50:228–33 CrossRef

13. Focke NK, Yogarajah M, Symms MR, et al. **Automated MR image classification in temporal lobe epilepsy.** *NeuroImage* 2012;59:356–62 CrossRef Medline

14. Chen S, Zhang J, Ruan X, et al. **Voxel-based morphometry analysis and machine learning based classification in pediatric mesial temporal lobe epilepsy with hippocampal sclerosis.** *Brain Imaging Behav* 2019. [Online ahead of print] CrossRef Medline

15. Ashburner J. **A fast diffeomorphic image registration algorithm.** *NeuroImage* 2007;38:95–13 CrossRef Medline

16. Kwan P, Arzimanoglou A, Berg AT, et al. **Definition of drug resistant epilepsy: consensus proposal by the ad hoc Task Force of the ILAE Commission on Therapeutic Strategies.** *Epilepsia* 2010;51:1069–77 CrossRef Medline

17. Riederer F, Marti M, Luechinger R, et al. **Grey matter changes associated with medication-overuse headache: correlations with disease related disability and anxiety.** *World J Biol Psychiatry* 2012;13:517–25 CrossRef Medline

18. Riederer F, Landmann G, Gantenbein AR, et al. **Nondermatomal somatosensory deficits in chronic pain are associated with cerebral grey matter changes.** *World J Biol Psychiatry* 2017;18:227–38 CrossRef

19. Wieser HG, Blume WT, Fish D, et al; Commission on Neurosurgery of the International League Against Epilepsy. **ILAE Commission Report. Proposal for a new classification of outcome with respect to epileptic seizures following epilepsy surgery.** *Epilepsia* 2001;42:282–86 Medline

20. Blumcke I, Thom M, Aronica E, et al. **International consensus classification of hippocampal sclerosis in temporal lobe epilepsy: a task force report from the ILAE Commission on Diagnostic Methods.** *Epilepsia* 2013;54:1315–29 CrossRef Medline

21. Nichols TE, Holmes AP. **Nonparametric permutation tests for functional neuroimaging: a primer with examples.** *Hum Brain Mapp* 2002;15:1–25 CrossRef Medline

22. Wagner J, Weber B, Urbach H, et al. **Morphometric MRI analysis improves detection of focal cortical dysplasia type II.** *Brain* 2011;134:2844–54 CrossRef Medline

23. Martin P, Winston GP, Bartlett P, et al. **Voxel-based magnetic resonance image postprocessing in epilepsy.** *Epilepsia* 2017;58:1653–64 CrossRef

24. McDonald CR, Hagler DJ Jr., Ahmadi ME, et al. **Subcortical and cerebellar atrophy in mesial temporal lobe epilepsy revealed by automatic segmentation.** *Epilepsy Res* 2008;79:130–38 CrossRef Medline

25. Sidhu MK, Duncan JS, Sander JW. **Neuroimaging in epilepsy.** *Curr Opin Neurol* 2018;31:371–78 CrossRef Medline

26. Azab M, Carone M, Ying SH, et al. **Mesial temporal sclerosis: accuracy of NeuroQuant versus neuroradiologist.** *AJNR Am J Neuroradiol* 2015;36:1400–06 CrossRef Medline

27. Farid N, Girard HM, Kemmotsu N, et al. **Temporal lobe epilepsy: quantitative MR volumetry in detection of hippocampal atrophy.** *Radiology* 2012;264:542–50 CrossRef Medline

28. Winston GP, Cardoso MJ, Williams EJ, et al. **Automated hippocampal segmentation in patients with epilepsy: available free online.** *Epilepsia* 2013;54:2166–73 CrossRef Medline

29. Coan AC, Kubota B, Bergo FP, et al. **3T MRI quantification of hippocampal volume and signal in mesial temporal lobe epilepsy improves detection of hippocampal sclerosis.** *AJNR Am J Neuroradiol* 2014;35:77–83 CrossRef Medline

30. Hammers A, Heckemann R, Koepp MJ, et al. **Automatic detection and quantification of hippocampal atrophy on MRI in temporal lobe epilepsy: a proof-of-principle study.** *NeuroImage* 2007;36:38–47 CrossRef Medline

31. Winston GP, Vos SB, Burdett JL, et al. **Automated T2 relaxometry of the hippocampus for temporal lobe epilepsy.** *Epilepsia* 2017;58:1645–52 CrossRef

32. Vasta R, Caligiuri ME, Labate A, et al. **3T magnetic resonance imaging simultaneous automated multimodal approach improves detection of ambiguous visual hippocampal sclerosis.** *Eur J Neurol* 2015;22:725–47 CrossRef Medline

33. Huppertz HJ, Wagner J, Weber B, et al. **Automated quantitative FLAIR analysis in hippocampal sclerosis.** *Epilepsy Res* 2011;97:146–56 CrossRef Medline

34. Good CD, Johnsrude I, Ashburner J, et al. **Cerebral asymmetry and the effects of sex and handedness on brain structure: a voxel-based morphometric analysis of 465 normal adult human brains.** *NeuroImage* 2001;14:685–700 CrossRef Medline

35. Good CD, Johnsrude IS, Ashburner J, et al. **A voxel-based morphometric study of ageing in 465 normal adult human brains.** *NeuroImage* 2001;14:21–36 CrossRef Medline

36. Huppertz HJ, Kroll-Seger J, Kloppel S, et al. **Intra- and interscanner variability of automated voxel-based volumetry based on a 3D probabilistic atlas of human cerebral structures.** *NeuroImage* 2010;49:2216–24 CrossRef Medline

37. Shinohara RT, the NAIMS Cooperative, Oh J, Nair G, et al. **Volumetric analysis from a harmonized multisite brain mri study of a single subject with multiple sclerosis.** *AJNR Am J Neuroradiol* 2017;38:1501–09 CrossRef Medline

38. Goubran M, Bernhardt BC, Cantor-Rivera D, et al. **In vivo MRI signatures of hippocampal subfield pathology in intractable epilepsy.** *Hum Brain Mapp* 2016;37:1103–19 CrossRef Medline

39. Jardim AP, Corso JT, Garcia MT, et al. **Hippocampal atrophy on MRI is predictive of histopathological patterns and surgical prognosis in mesial temporal lobe epilepsy with hippocampal sclerosis.** *Epilepsy Res* 2016;128:169–75 CrossRef Medline

40. Blumcke I, Pauli E, Clusmann H, et al. **A new clinico-pathological classification system for mesial temporal sclerosis.** *Acta Neuropathol* 2007;113:235–44 CrossRef Medline

41. Bonilha L, Edwards JC, Kinsman SL, et al. **Extrahippocampal gray matter loss and hippocampal deafferentation in patients with temporal lobe epilepsy.** *Epilepsia* 2010;51:519–28 CrossRef Medline

42. Bonilha L, Elm JJ, Edwards JC, et al. **How common is brain atrophy in patients with medial temporal lobe epilepsy?** *Epilepsia* 2010;51:1774–79 CrossRef Medline

43. Lee DH, Gao FQ, Rogers JM, et al. **MR in temporal lobe epilepsy: analysis with pathologic confirmation.** *AJNR Am J Neuroradiol* 1998;19:19–27 Medline