Multispectral Voxel-Based Morphometry of the Human Brain in Epilepsy

Dissertation

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

der Mathematisch-Naturwissenschaftlichen Fakultät

und

der Medizinischen Fakultät

der Eberhard-Karls-Universität Tübingen

vorgelegt

von

Raviteja, Kotikalapudi
aus Stadt Peddapuram, Land Indien

September - 2018

Tag der mündlichen Prüfung: 18/01/2019

Dekan der Math.-Nat. Fakultät: Prof. Dr. W. Rosenstiel

Dekan der Medizinischen Fakultät: Prof. Dr. I. B. Autenrieth

1. Berichterstatter: Prof. Dr.. Niels Fuke

2. Berichterstatter: Prof. Dr.. Uwe Klose

Prüfungskommission: Thomas Nägelle

Christoph Braun

Tag der mündlichen Prüfung: .....

Dekan der Math.-Nat. Fakultät: Prof. Dr. W. Rosenstiel

Dekan der Medizinischen Fakultät: Prof. Dr. I. B. Autenrieth

1. Berichterstatter: Prof. Dr.. …….....

2. Berichterstatter: Prof. Dr.. .....

Prüfungskommission: .....

.....

.....

.....

Erklärung / Declaration:

Ich erkläre, dass ich die zur Promotion eingereichte Arbeit mit dem Titel:

„Multispectral Voxel-Based Morphometry of the Human Brain in Epilepsy “

selbständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stellen als solche gekennzeichnet habe. Ich versichere an Eides statt, dass diese Angaben wahr sind und dass ich nichts verschwiegen habe. Mir ist bekannt, dass die falsche Abgabe einer Versicherung an Eides statt mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft wird.

I hereby declare that I have produced the work entitled “Multispectral Voxel-Based Morphometry of the Human Brain in Epilepsy”, submitted for the award of a doctorate, on my own (without external help), have used only the sources and aids indicated and have marked passages included from other works, whether verbatim or in content, as such. I swear upon oath that these statements are true and that I have not concealed anything. I am aware that making a false declaration under oath is punishable by a term of imprisonment of up to three years or by a fine.

Tübingen, den ....

Datum / Date Unterschrift /Signature

Contents

Abstract ... 1
1. General Introduction ... 2
1.1. Epilepsy and classification by ILAE Task Force ... 2
1.1.1. Seizure classification ... 2
1.1.2. Epilepsy classification ... 3
1.1.3. Epilepsy syndrome ... 3
1.2. Etiology ... 3
1.3. Focal epilepsy ... 4
1.4. Magnetic resonance imaging in focal epilepsy ... 5
1.4.1. Magnetic resonance imaging in cryptogenic focal epilepsy ... 6
1.5. Computational methods in structural MRI post-processing ... 6
1.6. Voxel-based morphometry (VBM) ... 7
1.6.1. Performance enhancement of MRI post-processing using VBM ... 8
1.7. Motivation ... 9
2. Summary ... 11
2.1. Study I ... 11
2.2. Study II ... 12
2.3. Study III ... 13
3. My contributions to scientific publications ... 14
3.1. Study I ... 15
3.2. Study II ... 15
3.3. Study III ... 16
4. Publications ... 18
4.1. Study I (published) ... 18
4.2. Study II (in press) ... 31
4.3. Study III (in preparation for submission) ... 43
5. Future works ... 79
6. Acknowledgment ... 80
7. References ... 81

Abstract

EPILEPSY is a burdensome neurological disorder, which affects people of all ages. However, the modern age of science, medicine and technology, has substantially helped our understanding of epilepsy. Of many medical examinations, magnetic resonance imaging (MRI) is one most important non-invasive ways to analyze and identify structural and functional brain abnormalities. Along with the advent of mathematical/statistical based computational models, brain morphometry can be very well studied and understood.

Voxel-based morphometry (VBM) is a well-established computational approach, which uses MRI images to detect structural differences between groups of subjects or single subjects against normal controls, the latter commonly used to detect lesions in focal epilepsy. The aim of my research is to validate existing VBM methods and explore strategies to improve its performance in detecting subtle lesions in patients, which can be potentially epileptogenic. Improvements in existing VBM can aid clinicians in diagnosing abnormal brain morphology. However, a good sensitivity (related to true positives) and a reasonable specificity (related to false positives) is paramount for its clinical usage.

My work in epilepsy is categorized into three major studies, for the dissertation;

i. Validation-comparison of multispectral segmentation based on 3D T1-, T2- and T2-FLAIR and its performance in detecting visible cortical lesions (Lindig et al., 2017).
ii. Optimization of multispectral voxel-based morphometry models in identifying epileptogenic findings in MRI-negative patients (Kotikalapudi et al., 2018).
iii. Applications of MP2RAGE sequences for multispectral voxel-based morphometry for enhancing lesion detection in focal epilepsy patients with conventional MRI images (Kotikalapudi et al., in preparation).

In conclusion, computational analysis of multispectral (-contrast) MRI sequences via voxel-based morphometry, shows promise of improving lesion detection, especially in focal epilepsy patients with previously normal conventional MRI.

1. General Introduction

The World Health Organization reports epilepsy as a chronic disorder of the brain, affecting approximately 50 million people (0.63%) worldwide (www.who.int). According to a meta-analysis study (of 222 studies), prevalence of active cases was reported at 6.38 per 1000 persons, lifetime prevalence at 7.6 per 1000 persons, whereas an annual cumulative incidence was 67.77 per 100,000 persons and incidence rate was 61.44 per 100,000 persons (Fiest et al., 2016). Epilepsy has been one of the most burdensome disorders of the brain and accounts to 0.70% of Disability-Adjusted Life Years, only after cerebrovascular disorders; 4.1%, headache; 0.97% and other neurological disorders; 0.72% (Chin and Vora, 2014). Therefore, understanding epilepsy and its classification is crucial, primarily for the clinicians' diagnosing patients. In addition, a clear classification contributes enormously to epilepsy research, development of new anti-epileptic therapies, and communication between epilepsy communities.

1.1. Epilepsy and classification by ILAE Task Force

The 2017 International League Against Epilepsy (ILAE), classifies epilepsy into three stages; seizure type, epilepsy type and epilepsy syndrome (Scheffer et al., 2017). In the present section, I attempt to briefly explain the latest classification system of ILAE, 2017.

1.1.1. Seizure classification

This classification framework begins with a systematic categorization of seizure type. An epileptic seizure is defined as "a transient occurrence of signs and/or symptoms due to abnormal excessive or synchronous neuronal activity in the brain" (Fisher et al., 2005). The seizures are classified into generalized epileptic seizures, focal epileptic seizures and unknown. Generalized seizures originate at some point within the brain, but rapidly spread in engaging bilateral neuronal networks, not confining to a single hemisphere. Focal seizures are abnormal neuronal activity in localized network(s) in the brain. Focal seizures include unifocal and multifocal seizures as well as seizures that are confined to one of the hemispheres. Every focal seizure is characterized by a consistent site, where it begins (onset). Unlike in the first two types, when it is not known whether a seizure is characterized by general or focal onset, it is labeled as unknown type. Unknown is a category stating that further clinical investigations have to be made to decide upon the type of seizure. A more detailed description can be obtained from ILAE, 2017 operational classification of seizure types (Fisher et al., 2017).

1.1.2. Epilepsy classification

Epilepsy is defined as "a disorder of the brain characterized by an enduring predisposition to generate epileptic seizures and by the neurobiologic, cognitive, psychological, and social consequences of this condition. The definition of epilepsy requires the occurrence of at least one epileptic seizure" (Fisher et al., 2005). Epilepsy type is broadly classified into four categories namely, generalized epilepsy, focal epilepsy, combined generalized and focal epilepsy and unknown category (Scheffer et al., 2017). For clinical diagnosis, generalized epilepsy is characterized by generalized onset of seizures, supported by generalized spike-wave activity on EEG findings. Focal epilepsies involve focal seizures, where diagnosis is performed on clinical grounds, supported by typical interictal EEG based focal epileptiform discharges. Combined generalized and focal epilepsy is a new category of epilepsy type, where patients show both; generalized and focal seizures. In these patients, diagnosis is performed on clinical grounds and supported by both generalized spike-wave and focal epileptiform discharges on an interictal EEG. Unknown epilepsy is used wherever the clinicians are not able to identify the patient with either focal or generalized epilepsies. This can be due to unavailability of sufficient information required for the clinical diagnosis and hence, further clinical investigations have to be performed.

1.1.3. Epilepsy syndrome

Features such as seizure onset, typical EEG attributes mapping to one or other types of epilepsies, anatomical and physiological features found through imaging modalities, genetic information, play a crucial role and aid in the clinical diagnosis of epilepsy. When these clusters of features occur together, it can be referred to as epilepsy syndrome (Scheffer et al., 2017). Epilepsy syndromes branch into more than 30 different categories and a detailed source of information is provided within the website, www.epilepsydiagnosis.org, developed by the ILAE Task Force. Childhood absence epilepsy, frontal lobe epilepsy, temporal lobe epilepsy, dravet syndrome, and west syndrome are some of the frequently used examples of epilepsy syndromes.

1.2. Etiology

At every classified stage i.e., seizure type, epilepsy type and epilepsy syndrome, the crucial goal of the clinician is in determining the underlying cause of the disorder, which is referred to as etiology. ILAE, 2017 broadly classifies etiology into six categories namely, structural, genetic, metabolic, immune, infectious and unknown (Scheffer et al., 2017). The term 'structural etiology' is used, when an underlying structural abnormality is the most likely cause of the epilepsy (Berg et al., 2010), e.g. mesial temporal lobe epilepsy with hippocampal sclerosis or epilepsy due to malformations of cortical development are archetypes for a structural etiology. Structural abnormalities may be acquired through stroke, injury, trauma, infection and/or caused by genetic mutations. In case of tuberous sclerosis complex (TSC), while malformations in the cortex through cortical tubers form the structural etiology,

mutations of TSC1/2 gene associate with genetic etiology (Han and Sahin, 2011). The concept of 'genetic etiology' stems from a known or presumes gene mutation, which increases the risks for epileptic seizures. In benign familial neonatal epilepsy, studies have shown a 91% association of a molecular cause with mostly mutations of KCNQ2 and also SCN2A gene with onset seizures in 36 families (Grinton et al., 2015). Seizures can also trigger from a variety of metabolic dysfunctions. For example, in patients with hyponatraemia, seizures can occur as an effect from reduced serum levels, mostly when below 115mmol/L (Delanty et al., 1998). 'Infection etiology' is the most common cause of epilepsy worldwide, where infection conditions such as acute bacterial meningitis, intracranial abscesses, central nervous system tuberculosis, and cerebral malaria have been linked with a high prevalence of epilepsy (Vezzani et al., 2016). (Auto-) immune system disorder leading to a higher risk of epilepsy is conceptualized with 'Immune etiology'. For instance, nearly 10-20% of patients with systemic lupus erythematosus (SLE) suffer epileptic seizures at some point in their disorder induced life, due to immune-mediated neuronal damages (Aarli, 2000). Unknown etiology is used wherever the underlying cause is yet to be identified.

1.3. Focal epilepsy

Focal epilepsy is one of the common types of epilepsies, where seizures originate from a localized region in the brain. The most important drug-resistant focal epilepsy syndromes include neocortical epilepsy, associated with focal cortical dysplasia (FCD) and temporal lobe epilepsy with mesial temporal sclerosis (MTS). FCD is a subgroup of malformation of cortical development (MCD), the latter including a variety of developmental disorders that serve as a common cause for epilepsy (Barkovich et al., 2012). Focal dysplasia of the cortex was first described in 1970s, based on a series of surgically resected epilepsy specimens (Taylor et al., 1971). FCDs are frequently depicted as common etiologies causing epilepsy in child and adult populations (Blumcke et al., 2011). Based on the histopathology findings, ILAE 2011 have predominantly classified FCD into Type I, II and III (Blumcke et al., 2011). In brief, various histopathological findings include abnormal radial/tangential lamination of the neocortex (FCD Type I), disrupted cortical architecture and cytoarchitectural abnormalities (FCD Type II), dysplasias associated with adjacent causes of epilepsy (FCD Type III) including hippocampal sclerosis, tumors (e.g., glangliogliomа), vascular malformations (e.g., cavernomas), and/or with lesions acquired during early life. Other MCD conditions include polymicrogyria, schizencephaly, periventricular nodular hypertropia and hemimegalencephaly (Barkovich et al., 2012). In the same study, genetic classification of MCDs has also been covered in detail. Mesial temporal lobe epilepsy with hippocampal sclerosis (mTLE-HS) is also one of the most common types of focal epilepsy of the temporal lobe, often not responding to antiepileptic drugs (Téllez-Zenteno and Hernández-Ronquillo, 2012). ILAE commission in 2004 reported in detail, a group of pathological conditions serving as minimal criteria for the diagnosis of TLE with HS (Wieser, 2004). In brief, these pathological conditions include neuronal loss and gliosis to variable degrees in hippocampal areas, changes in the glial cellular

organization, extrahippocampal pathology depicted at structures including amygdala and other mesial temporal lobe and/or white matter structures of temporal lobe.

1.4. Magnetic resonance imaging in focal epilepsy

For over three decades, magnetic resonance imaging (MRI) has been consistently used in epilepsy, primarily for the detection of underlying structural abnormalities. MRI is widely available and often the primary non-invasive imaging modality of choice in epilepsy (Duncan, 2002). The main goal of MRI is to detect underlying pathologies including tumors, vascular lesions and also aid clinicians in formulating syndromes and etiologies (Berkovic et al., 1998). As a routine clinical practice, MRI image acquisition protocols include whole brain T1-weighted and T2-weighted sequences, with minimum slice thickness, which can be best supported by the scanner. Though there are several such guidelines, an official standard protocol has not been presented so far. However, it is recommended to have T1 volumetric acquisition (3D) with inversion recovery at an isotropic resolution of 1mm-1.5mm, facilitating image reconstruction in any planar direction (Cendes et al., 2016). Along with T1, T2 and T2-weighted fluid attenuation inversion recovery (FLAIR) sequences (<3mm slice thickness) with at least 2 cut-plane orientations; axial and coronal, are also acquired in hippocampal angulation (Spencer, 2014). Using these sequences, already a wide range of focal epileptogenic lesions can be visually determined. For instance, qualitative assessment on T1 weighted MRI identified cortical thickening in 50-92% cases (FCD Type IIb), blurring of gray-white matter junction in 60-80% case, while a hyperintense T_2 signal revealed in 46-92% of FCD lesions (Bernasconi et al., 2011; Blumcke et al., 2011). In patients with mTLE-HS, abnormal hippocampal atrophy with an increased T2 (or hypointense T1) signal can be qualitatively assessed (Cendes et al., 2016). These are certain distinct features, neuroradiologists look for in MRI images, as image biomarkers for focal epilepsy. Hence, MRI sequences, especially T1, T2 and FLAIR are extremely helpful in the diagnosis of focal epilepsy. Apart from these sequences, diffusion weighted imaging (DWI), hemosiderin and calcification sensitive sequences, and contrast-enhanced T1 sequences are also useful (Wellmer et al., 2013). Also development of new sequences such as magnetization prepared 2 rapid gradient echo (MP2RAGE) provide better tissue contrasts, in comparison to T1-weighted sequences (Marques et al., 2010). However, MP2RAGE diagnostic yield in focal epilepsy is yet to be validated. Apart from MRI, studies have also presented in detail, other popular imaging methods in epilepsy including PET; positron emission tomography, SPECT; single-photon emission computed tomography, MEG; magnetic encephalography, DTI; diffusion tensor imaging; fMRI and EEG; function imaging with electroencephalography (Cendes et al., 2016; Duncan, 1997; Kini et al., 2016).

1.4.1. Magnetic resonance imaging in cryptogenic focal epilepsy

In drug-resistant focal epilepsy, often surgical resection brings the best outcomes of a seizure free life (Wiebe et al., 2001). Better detection and localization of lesion on a preoperative MRI not only improves presurgical diagnosis, but can also increases seizure-freedom post-surgery. However, often MRI is unremarkable (or negative) in approximately half of focal epilepsy cases (McGonigal et al., 2007). Further large cohort studies have shown that patients with a non-lesional MRI suffer from poor surgical outcomes (Tellez-Zenteno et al., 2010). But failure to identify lesions on an MRI, does not necessarily determine the patients to be non-lesional. For example, focal cortical dysplasia has been reported to be the most common pathology (43%-58% cases) in unremarkable MRI patients (Bien et al., 2009; Wang et al., 2013). It should be appreciated that lesion detection significantly improved with advancements in MRI hardware systems from 1.5T to 3T (Knake et al., 2005). Yet, 50-80% FCD lesions escape visual detection (Besson et al., 2008). Tassi et al., reported 34% histopathologically proven FCDs were visually detected on MRI, after re-evaluation, from previously MRI negative scans (Tassi et al., 2002). This shows that, with advanced MRI technology, it is also important to qualitatively assess these lesions through expert visual detection. Studies also demonstrate varied visual interpretation rates from non-expert to expert (39% to 50%) in detecting structural lesions on MRI (Von Oertzen et al., 2002). Hence, the task of visually analyzing lesions presents us with a considerable variability among clinicians. Therefore, technical advancements in MRI are important, but may not always be sufficient in improving lesion detection. With the advent of computational methods for MRI post-processing, finding its potential in detecting lesions in this more challenging MRI negative cohort is crucial and highly valuable. Automated methods can aid in improving visual detection of lesions, thus increasing the diagnostic yield of present MRI systems in epilepsy.

1.5. Computational methods in structural MRI post-processing

Computational methods are gaining rapid prominence in identifying structural malformations in epilepsy. Most of these methods are developed to identify/detect pattern abnormalities such as gray-white matter junction blurring, variable cortical thickening, sulcus and gyral patterns, and regions with hyperintensities (Kini et al., 2016). In epilepsy, over the last decade, voxel based morphomtery (VBM), has been one of the most widely used methods in detecting such structural malformations, especially abnormal gray matter. In brief, VBM uses a systematic approach, by which voxel-wise comparison of concentrations or volumes between two groups of brains can be facilitated (Ashburner and Friston, 2000). The characteristic blurring of GM-WM junctions in FCDs has been most consistently identified among several methods, using MAP technique which generates junction maps (Huppertz et al., 2005). Junction maps are primarily used to isolate the voxels, which fall in between GM-WM intensities. Surface based morphometry (SBM) is also a common group of computational techniques used primarily to perform reconstruction and analysis of cortical surface. Most common surface based analysis (SBA) method in use is Freesurfer. It a very useful method to perform cortical surface analysis and investigate

architectural attributes including cortical thickness, cortical and sub cortical volumes, and curvature properties (Dale et al., 1999).

While these methods are predominantly used on high resolution T1-weighted images, computational methods can also be implemented on other MRI sequences. Apart from standard T1, voxel based intensity analysis methods also include the use of sequences such as T2, and FLAIR. in focal epilepsy. For instance, Focke et al., describes a novel voxel based intensity analysis similar to standard VBM, but using intensity normalized FLAIR images (Focke et al., 2008). House et al., describes the use junction maps, but derived from T1-weighted images (House et al., 2013). Methods including DTI, MR-relaxometry, magnetization transfer imaging (MTI), support vector machines (SVM) also find their application in epilepsy and have been covered in detail, in epilepsy reviews of the past (Kini et al., 2016; Martin et al., 2015).

1.6. Voxel-based morphometry (VBM)

VBM follows a systematic pipeline comprising several inter-dependent steps, to achieve visualization of structural differences, as the final output. Information of each step, its significance and method validation are presented in detail in the articles on VBM methods (Ashburner and Friston, 2000; Good et al., 2001). In brief, the processing steps include, 1. tissue segmentation; segmenting MRI images primarily into three tissue classes namely, gray matter (GM), white matter (WM), cerebrospinal fluid (CSF) using prior tissue probability information, 2. normalization; representing the classified tissues in a stereostatic space (common space) via tissue deformation, 3. smoothing; eliminating or attenuating tissue inhomogeneities and conforming images closer to normal distribution and 4. statistical comparisons; facilitating statistical tests to elevate significant differences. This method is available via a toolbox with statistical parametric mapping (SPM; www.fil.ion.ucl.ac.uk). VBM is primarily implemented in focal epilepsy, based on T1 images, in detecting visually identified MCDs on MRI (Bonilha et al., 2006a; Colliot et al., 2006). Implementing VBM has resulted in detecting a vast majority of MRI-visible lesions.

For lesional MRI cases, T1 VBM has yielded appreciable sensitivity in the range of 38%- 95% (Martin et al., 2015). Voxel based analysis of whole brain 3D FLAIR reported a sensitivity of 88% (22/25 cases) in detecting neuroradiologically proven FCDs. A voxel based analysis of junction maps based on T1 showed a sensitivity of 84% in detecting proven FCD lesions (House et al., 2013; Huppertz et al., 2005). However, while considering the more challenging cohort of unremarkable (negative/non-lesional) MRI cases, voxel-based analysis has suffered from low detection. For instance, in a cohort of 13 TLE-HS patients, T1 VBM could not detect HS or cortical neuronal loss in these epilepsy patients (Eriksson et al., 2009). In one of the studies, T1 VBM could not detect any abnormalities in MRI negative children (14 cases) with cryptogenic focal epilepsy, while VBM of FLAIR could only detect 14% (2 cases),

correlating with other clinical data on epileptogenic zone (Riney et al., 2012). In another study of 70 patients with refractory focal epilepsy, FLAIR VBM findings coincided with scalp video-EEG telemetry in only 11.4% cases (Focke et al., 2009). Thereby, the overall detection rate of voxel-based analyses in MRI-negative cases is considerably low in comparison to lesional MRI cases, while using one of the sequences, either T1, T2 or FLAIR.

We have seen that, VBM is a promising computational method and has been established in detecting structural abnormalities in known lesional MRI cases. However, its low detection rate in MRI-negative cases requires the need for systematic assessment, validation and methodological interventions in VBM processing steps/parameters. This would help in enhancing the diagnostic yield of VBM in 'previously missed lesions via visual analysis of MRI images'. Hence, a systematic comparison of VBM, which includes comparison of MRI sequences and method parameters such as segmentation, smoothing, and statistical thresholds, is very essential.

1.6.1. Performance enhancement of MRI post-processing using VBM

Unified segmentation was introduced in 2005 and has been incorporated in SPM software and subsequent versions. In brief, unified segmentation is a generative model which includes three processes namely, image registration, tissue classification and bias correction (Ashburner and Friston, 2005). Incorporating this segmentation method, multimodal segmentation can be performed in SPM. It was first incorporated in SPM8 via 'new segment'. In theory, multimodal segmentation involves simultaneous inclusion of two or more co-registered image contrasts of the same subject to perform tissue classification. Recently, multispectral segmentation with addition of FLAIR contrast, has proved to be advantageous over T1 (Viviani et al., 2017b). The advantages include a better differentiation between GM and dura, and exclusion of vessels from GM. However, a detailed performance analysis based on T1, T2 and FLAIR for segmenting anatomical regions into GM, WM, and CSF and the best performing combination of these tissue contrasts is still unknown. Further, the applications of multimodal/multiple contrast based models in focal epilepsy are yet to be assessed. Also, with newer sequences such as MP2RAGE, promising a better differentiation of GM-WM boundaries in comparison to existing (Marques et al., 2010) sequences, is yet to be validated through multimodal tissue classification approach, in focal epilepsy.

Second, an important challenge in VBM for effective lesion detection is the choice of several parameters, of which two crucial parameters are smoothing and statistical threshold. The size of the smoothing kernel determines the number of voxels, which are averaged at each point. By conforming the image intensities closer to a Gaussian distribution, it contributes to reducing inter-subject variability (Good et al., 2001). While smoothing increases the sensitivity of the method by reducing inter-subject variability, excess smoothing can diminish subtle findings, which may be crucial in certain pathologies.

Further, while analyzing especially MRI-negative cases for structural abnormalities, these subtle changes, which can be overlooked in visual assessment, are very sensitive to the size of Gaussian kernel applied. In focal epilepsy, studies have used different smoothing Gaussian kernels, and an optimized smoothing strategy is still unknown (Huppertz et al., 2005; Keller and Roberts, 2008).

Third, the choice of statistical threshold has a potential impact on VBM findings, and has remained empirical. Stricter threshold of $p<0.05$ family wise error correction (FWE) and liberal threshold of $p<0.01$ uncorrected, have drastically affected concordant findings in patients and false positive findings in controls (Martin et al., 2017). The choice of statistical thresholds does not necessarily affect a clinicians' perspective towards lesions, which are clearly visible on MRI images. But, misinterpretation or over interpretation of VBM findings is a setback while assessing an MRI-negative cohort, questioning the reliability of the method. Studies in the past have used variable statistical thresholds and a systematic validation of these cut-offs are not mentioned (Bonilha et al., 2006a; Colliot et al., 2006; Huppertz et al., 2005; Keller and Roberts, 2008; Martin et al., 2017). Therefore, it is important to systematically compare the power of statistical thresholds, for striking a better trade-off between true positive findings in patients (sensitivity) and false-positive findings in controls (specificity).

Hence, it is important to incorporate multispectral VBM methods based on a combination of different MR contrasts (sequences/images) and study their performance in epilepsy, but at variable spatial and statistical parameters.

1.7. Motivation

The major goal of research is to deepen our understanding of an issue, which is considered to be of significant relevance to our lives. Epilepsy, being one of the most burdensome neurological conditions, has affected the functional performance of people suffering from it. However, with the advancements in science and technology, we are highly motivated in improving their lives. Also, this will help us contribute to the betterment of the society, we live in. Machines and their computational methods are slowly gaining applicability in medicine. Despite, they always provide to us, a plethora of scope for improvement to fill in the existential gaps between artificial intelligence and human intelligence. This scope serves as my philosophical motivation for research.

Stating that, "automated computational methods help clinicians in improving their decisions in epilepsy", is not an understatement. But, especially in the most popular methods; VBM, there still lies the task of a systematic assessment and validation of the techniques in epilepsy. I feel that the true potential of computational methods in aiding clinicians for diagnostic decisions is yet to be explored. This entire research aims at providing reliable automated results to clinicians, which can be referred to, while taking important presurgical diagnosis decisions. Especially, in focal epilepsy patients with

conventional normal MRI, through a series of research studies we aim at improving the performance of existing VBM methodologies. These motivations will not only increase the conversion of MRI-negative to MRI-positive patients, but also improve our understanding and diagnosis in epilepsy.

2. Summary

My research contributions for this dissertation have been summarized into three publications (status being published (1), in press (1) and under preparation for submission (1)). The three studies namely Lindig et al., 2017 (published), Kotikalapudi et al, 2018 (in press) and Kotikalapudi et al. (under preparation for submission), systematically compare existing T1-weighted VBM (T1 VBM) with newer multispectral VBM variants based on T2-weighted sequences (T2 and FLAIR) and newer sequences mainly MP2RAGE. Most importantly, the systematic comparisons are also facilitated by the optimization of crucial spatial and statistical parameters in the VBM methods. Eventually, these comparisons lead to findings the diagnostic performance of the VBM variants in detecting structural brain abnormalities in patients diagnosed with focal epilepsy.

2.1. Study I

In the first study (Lindig et al., 2017), we have systematically validated brain tissue segmentation approaches, which is the primary start point for voxel-based analysis (VBM). This was achieved by conducting a systematic validation and comparison of different segmentation approaches namely single-channel and multispectral (-channel,-contrast, modal) segmentations using statistical parametric mapping (SPM12, https://www.fil.ion.ucl.ac.uk/). For this study, we recruited 92 healthy volunteers (controls), 5 patients with TLE-LHS (temporal lobe epilepsy with left hippocampal sclerosis) and 7 patients with malformation of cortical development (MCD). 3D T1-weighted, T2-weighted and T2-weighted FLAIR images were acquired at 3T Siemens MRI for all the subjects using the protocol mentioned in the published study (Lindig et al., 2017), with an isotopic resolution of $0.9mm^3$. Unified segmentation method (Ashburner and Friston, 2005) was implemented for single-channel T1 and multispectral combinations of T1+T2, T1+FLAIR, T2+FLAIR and T1+T2+FLAIR. Next, the segmentation approaches were validated for their effects on gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) tissue classification with respect to population age in the control cohort. Later, VBM was performed for all the 5 segmentation combinations and systematic comparisons were drawn commenting on quality of tissue segmentation. Finally, single-channel and multispectral VBM variants (models) were tested for their ability to identify epileptogenic lesions in a focal epilepsy patient population with visually identified lesional MRI cohort. All the computational findings were reviewed by two expert neuroradiologists with approximately 10 years of experience.

Upon facilitating statistical comparisons between segmentation approaches to validate the group effects, we made various observations. Most importantly, we found that single-channel T1 segmentation misclassifies/overestimates dura (meninges) and blood vessels (mainly sagittal, transverse and sigmoid sinuses) as GM. This problem arises due to the isointensity of these tissues with GM, hence segmented

as GM. This problem is most readily solved with the addition of FLAIR channel to T1 i.e., T1+FLAIR. Primarily reason for addressing the overestimation of GM is the fact that vessels and dura are suppressed/attenuated by applying FLAIR sequences, and this effect is readily translated in multispectral segmentation approach. In the immediately mentioned study, they had similar findings addressing overestimation of vessels and dura with the application of MPRAGE+FLAIR-based multispectral segmentation. Also, for both single channel and multispectral segmentation approaches, we found that absolute volumes of GM decreased linearly with an increase in age. For this linear regression model, multispectral segmentation combination of T1+T2+FLAIR demonstrated the best goodness of fit. Similar results were also obtained for T1+FLAIR and T2+FLAIR. Hence, a comparison of different segmentation combinations for GM tissue classification not only helps in validating the models, but also gives us information regarding the model with the best fit, to be a superior segmentation approach. As a second validation step, we found that multispectral FLAIR based VBM variants, especially T1+FLAIR and T2+FLAIR showed improved lesion detection in patients with PM and MCD. Also, in the TLE-LHS group, multispectral T1+T2 performed superior to existing T1 VBM. Please refer to our scientific article Lindig et al., 2017, for further details from this study.

In brief, performance of multispectral VBM variants in comparison to existing single-channel T1 VBM in:

1. Better classification brain tissue into GM, WM and CSF,
2. Generating superior goodness of fit for the effects of tissue volumes with age,
3. Improved detection of lesions in patients with MCD/TLE-HS,

shows promise of testing multispectral VBM variants to detect subtle/occult malformations in the most challenging cohorts of MRI-negative focal epilepsy. This serves as the most important question, which was addressed in Study II.

2.2. Study II

In this study (Kotikalapudi et al., 2018), we have developed a novel approach for conducting a systematic comparison across single-channel T1 and multispectral VBM variants T1+T2, T1+FLAIR and T1+T2+FLAIR in detecting subtle epileptogenic brain morphologies in previously MRI-negative focal epilepsy cohort. For this study, we recruited 62 healthy controls and 13 patients diagnosed with focal epilepsy with a lobar clinical hypothesis of seizure onset, but normal (/negative) conventional MRI, upon visual review. VBM analysis for identifying gray matter concentration and volume changes were performed for all 4 variants. To address specificity of the VBM findings, each control was compared against the rest of the controls (control Cx in comparison versus 61Cx) at multiple smoothing levels, 4mm to 16mm and increasing statistical cutoffs 2.5 to 6. In a similar way, each patient (Px) was compared against all the controls (Px versus 62Cx) to test for the concordance of the findings with the

lobar hypothesis. Similarly, discordant findings (findings deferring from clinical hypothesis) were also analyzed in the patient population. The performance of VBM variants at each smoothing level was quantified by implementing receiver operating characteristic curve (ROC) analysis. All the automated findings were visually inspected by an expert neuroradiologist with 10 years of experience, and given clinical ratings on whether they held clinical relevance or not. For more details, also please refer to Kotikalapudi et al., 2018.

We found that at higher smoothing levels, 12mm and liberal statistical cutoffs, 3.7, almost all VBM variants showed a balance between concordant findings, specificity and concordant to discordant ratio in patients. Most importantly, in automated VBM analysis and after visual review of VBM findings, we found that T1+FLAIR was superior amongst all VBM variants and when compared to single-channel T1 VBM. One of the main reasons behind this effect was the superior segmentation of potentially epileptogenic lesions as GM, by the addition of mainly FLAIR channel to existing T1. The performance of FLAIR based voxel analysis is in line with several studies in MRI negative focal epilepsy (Focke et al., 2009, Riney et al., 2012). Most importantly, we found that superior tissue segmentation for multispectral VBM variants (Study I-section 2.1) was responsible for the better performance of such models (T1+FLAIR) in detecting subtle abnormalities in MRI-negative focal epilepsy.

In the past, limited literatures have reviewed systematic comparison and optimization of VBM methods in improving lesion detection in focal epilepsy (Martin et al., 2017). Our study can serve as a basic strategy approach in assessing the reliability of VBM findings, when presented with a reasonable specificity. Without a good balance between sensitivity and specificity, these findings can be easily mis-/over-interpreted, when talking of subtle morphological differences in brain tissues. Also, through this study, we show that multispectral VBM can be a potential upgrade over existing T1 VBM methods. However, these findings should be considered in the presurgical evaluation of epilepsy only after a careful review by clinicians.

T1-based multispectral VBM variants are valuable in improving clinically relevant information in focal epilepsy. But, especially at higher static magnetic field strengths ≥3T, 7T and 9.4T, T1 and T2-weighted sequences suffer from radio-frequency field inhomogeneities. In such critical scenarios, careful evaluation of newer VBM models utilizing self-bias corrected sequences mainly MP2RAGE (Magnetization Prepared 2 Rapid Gradient Echo) becomes highly important. Exploration of MP2RAGE-based VBM approach in focal epilepsy, by implementing the optimized VBM model from this study (II) serves as the main aim for study III.

2.3. Study III

In Study II we optimized VBM analysis for single-channel as well as multispectral combinations, mainly dealing with T1-, T2-, and FLAIR-weighted images. To continue further, we assessed the performance of self-bias correcting sequence, MP2RAGE in brain tissue classification. Next,

multispectral VBM variants of MP2RAGE, such as MP2+INV1 and MP2+INV2 were also tested and their effect in detecting subtle/hidden epileptogenic abnormalities was quantified. Comparisons were also drawn with existing T1 and recent T1+FLAIR VBM methods in focal epilepsy. Further details especially on MP2RAGE based combinations and their performance in focal epilepsy, have been out in details in Kotikalapudi et al., which is in preparation for submission.

For this study, we recruited 31 healthy volunteers and 21 patients with focal epilepsy (5 with visible lesions and 16 MRI-negative). Implementing the methodological designs from study I and study II, we systematically compared the tissue segmentations and performance of MP2RAGAE VBM variants with existing T1 and T1+FLAIR methods in focal epilepsy. We found several advantages for the MP2RAGE based combinations such as MP2+INV1 and MP2+INV2. These immediately mentioned combinations can be used especially to enhance tissue segmentation process in cortical and subcortical GM areas. Also while comparing the performances of VBM variants, we found a superiority in MP2RAGE based VBM variants in detecting subtle epileptogenic lesions, definitely in comparison to most commonly used single-channel T1 VBM.

In conclusion, we find that MP2RAGE-based multispectral VBM variants are highly useful in lesion detection at 3T. They also hold additional advantage of self-bias correction at higher static magnetic field-strengths ≥7T/9.4T.

3. My contributions to scientific publications

In this section, my contributions to this thesis, in terms of scientific articles have been provided.

3.1. Study I

Study I refers to:

"Evaluation of multimodal segmentation based on 3D T1-, T2- and FLAIR-weighted images – the difficulty of choosing"

Tobias Lindig, **Raviteja Kotikalapudi**, Daniel Schweikardt, Pascal Martin, Friedemann Bender, Uwe Klose, Ulrike Ernemann, Niels K. Focke, Benjamin Bender

Author contributions

TL, BB and NKF helped in the conceptualization of this study. Data acquisition and management was done by TL, DS, PM, FB and BB. Original draft was written by TL and BB. Manuscript revisions including crucial commentary was provided by TL, RK, DS, PM, FB, UK, UE, NKF and BB. Manuscript editing and revision was done by TL, RK (analysis, tables and figures) and BB. Project administration was done by UE and BB. Correspondence of the manuscript with the publishing journal was done by NKF.

My contributions

I, Raviteja Kotikalapudi (RK) developed the computational methodological programming stream and conducted formal analysis of the data. I produced the 'results' that form the basis for this study and made the 'table and figures' for the scientific article. I also contributed to the writing, for the 'methods' section.

Funding

The financial support for this study was received from the Medical Faculty of the University of Tübingen, Germany (grant number AKF 321-0-1).

Acknowledgement

The authors want to thank Merim Bilalic and Matthew Bladen who proof-read the manuscript and gave valuable suggestions.

3.2. Study II

Study II refers to:

"Systematic assessment of multispectral voxel-based morphometry in MRI previously negative for focal epilepsy"

Raviteja Kotikalapudi, Pascal Martin, Justus Marquetand, Benjamin Bender, Niels K. Focke (In press)

Author contributions

RK and NKF conceptualized this study. Data acquisition and management was done by RK, PM, JM and BB. Original draft was written by RK. Data validation and visualization was performed by RK, PM, BB and NKF. Manuscript revisions inclusions crucial comments were given by PM, JM, TL and BB. Manuscript editing and revision was done by RK and NKF. Project administration was done by NKF. Correspondence of the manuscript with the publishing journal was done by RK.

Funding

The financial support for this study was received from the Faculty of medicine, University of Tübingen (321-0-1) and the grant; Deutsche Forschungsgemeinschaft bearing grant number FO750/5-1.

3.3. Study III

Study III refers to:

"Multispectral voxel based morphometry of MP2RAGE sequences and comparison with existing T1-weighted and T2-weighted sequences in focal epilepsy"

Raviteja Kotikalapudi*, Pascal Martin*, Michael Erb, Justus Marquetand, Benjamin Bender, Niels K. Focke (In preparation for scientific article in journal NeuroImage: Clinical)

*Equally contributing

Author contributions

RK and NKF conceptualized this study. Data acquisition and management was done by RK, PM, JM and ME. Original draft was written by RK. Data validation and visualization was performed by RK, PM, BB and NKF. Manuscript revisions which included crucial suggestions were given by PM, ME, JM and BB. Manuscript editing and revision was done by RK and NKF. Project administration was done by NKF. Correspondence of the manuscript with the publishing journal was done by NKF.

Funding

The financial support for this study was received from the grant; Deutsche Forschungsgemeinschaft bearing grant number FO750/5-1.

4. Publications

4.1. Study I (published)

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Evaluation of multimodal segmentation based on 3D T1-, T2- and FLAIR-weighted images – the difficulty of choosing

Tobias Lindig[a], Raviteja Kotikalapudi[a,b], Daniel Schweikardt[a,1], Pascal Martin[b], Friedemann Bender[a,b], Uwe Klose[a], Ulrike Ernemann[a], Niels K. Focke[b,*], Benjamin Bender[a]

[a] Dept. of Diagnostic and Interventional Neuroradiology, University Hospital Tübingen, Hoppe-Seyler-Str. 3, 72076 Tübingen, Germany
[b] Neurology and Epileptology, Hertie Institute for Clinical Brain Research, University Hospital Tübingen, Hoppe-Seyler-Str. 3, 72076 Tübingen, Germany

ARTICLE INFO

Keywords:
Automatic lesion detection
Epilepsy
Atrophy
Statistical parametric mapping

ABSTRACT

Voxel-based morphometry is still mainly based on T1-weighted MRI scans. Misclassification of vessels and dura mater as gray matter has been previously reported. Goal of the present work was to evaluate the effect of multimodal segmentation methods available in SPM12, and their influence on identification of age related atrophy and lesion detection in epilepsy patients. 3D T1-, T2- and FLAIR-images of 77 healthy adults (mean age 35.8 years, 19–66 years, 45 females), 7 patients with malformation of cortical development (MCD) (mean age 28.1 years,19–40 years, 3 females), and 5 patients with left hippocampal sclerosis (LHS) (mean age 49.0 years, 25–67 years, 3 females) from a 3 T scanner were evaluated. Segmentation based on T1-only, T1+T2, T1+FLAIR, T2+FLAIR, and T1+T2+FLAIR were compared in the healthy subjects. Clinical VBM results based on the different segmentation approaches for MCD and for LHS were compared. T1-only segmentation overestimated total intracranial volume by about 80 ml compared to the other segmentation methods. This was due to misclassification of dura mater and vessels as GM and CSF. Significant differences were found for several anatomical regions: the occipital lobe, the basal ganglia/thalamus, the pre- and postcentral gyrus, the cerebellum, and the brainstem. None of the segmentation methods yielded completely satisfying results for the basal ganglia/thalamus and the brainstem. The best correlation with age could be found for the multimodal T1+T2+FLAIR segmentation. Highest T-scores for identification of LHS were found for T1+T2 segmentation, while highest T-scores for MCD were dependent on lesion and anatomical location. Multimodal segmentation is superior to T1-only segmentation and reduces the misclassification of dura mater and vessels as GM and CSF. Depending on the anatomical region and the pathology of interest (atrophy, lesion detection, etc.), different combinations of T1, T2 and FLAIR yield optimal results.

Introduction

Over the last decade, voxel-based morphometry (VBM) (Ashburner and Friston, 2005) has been mainly based on T1-weighted magnetic resonance images to identify volume changes within normal aging and regional brain atrophy in a variety of neurological and psychiatric diseases. The misclassification of brain structures with poor gray and white matter contrast like the basal ganglia, thalamus and brainstem (Good et al., 2001) or structures with similar intensities on T1-weighted images such as dura versus gray matter (van der Kouwe et al., 2008; Viviani et al., 2016) or gray matter versus vessels (Viviani et al., 2016) is an ongoing and still unsolved issue. In theory, a multimodal segmentation approach that combines images with different contrast for each tissue type should perform better than an approach based on T1-weighted images only. This has already been shown for the brainstem by Lambert et al. based on images with a T1, MT and PD contrast (Lambert et al., 2013). It has also been shown that standard segmentation based on T1-weighted images overestimates gray matter volume by wrongly assigning vessels and dura to this compartment and that the combination of T1 and FLAIR images

Abbreviations: CSF –, cerebrospinal fluid; FCD –, focal cortical dysplasia; FWE –, family-wise error rate; GM –, gray matter; LHS –, left hippocampal sclerosis; MCD –, malformation of cortical development; SD –, standard deviation; SE –, standard error; TIV –, total intracranial volume; TLE –, temporal lobe epilepsy; VBM –, voxel-based morphometry; WM –, white matter.

* Corresponding author.
E-mail addresses: tobias.lindig@med.uni-tuebingen.de (T. Lindig), raviteja.kotikalapudi@gmail.com (R. Kotikalapudi), daniel.schweikardt@gmail.com (D. Schweikardt), pascal.martin@med.uni-tuebingen.de (P. Martin), friedemann.bender@gmx.de (F. Bender), uwe.klose@med.uni-tuebingen.de (U. Klose), ulrike.ernemann@med.uni-tuebingen.de (U. Ernemann), niels.focke@uni-tuebingen.de (N.K. Focke), benjamin.bender@med.uni-tuebingen.de (B. Bender).
[1] Present address: Institute of Diagnostic and Interventional Radiology, RKH Klinikum Ludwigsburg, Ludwigsburg, Germany.

http://dx.doi.org/10.1016/j.neuroimage.2017.02.016
Received 26 October 2016; Accepted 6 February 2017

Please cite this article as: Lindig, T., NeuroImage (2017), http://dx.doi.org/10.1016/j.neuroimage.2017.02.016

results in a better segmentation (Viviani et al., 2016). Some authors suggest that in pathologic tissue T2-weighted VBM should perform better to identify tissue pathology (Diaz-de-Grenu et al., 2011). Starting with SPM8 ("New Segment") the possibility of multimodal segmentation has been added to Statistical Parametric Mapping (Wellcome Trust Centre, 2017) and the segmentation approach has become the routine segmentation algorithm in SPM12.

A detailed analysis of the performance of this multimodal approach of tissue classification in gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) and its dependence on anatomic locations is lacking. Moreover, it is unknown which combination of T1, T2 or FLAIR is preferable. Therefore, the aim of the present study is to assess the difference of SPM12 multi-parametric tissue segmentation based on combinations of high resolution T1-, T2- and FLAIR weighted images in comparison to the standard T1 based segmentation, and its effect on assessment of age related changes and on automatic lesion detection in patients with lesional focal epilepsy.

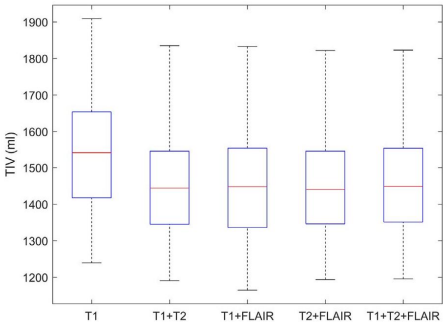


Fig. 1. Boxplot of the total intracranial volume (TIV) for the different segmentation methods.

Methods

Healthy subjects

The cross-sectional prospective study was approved by the local institutional review board, and all healthy subjects gave written informed consent prior to the examination. Exclusion criteria were known present or past neurological and psychiatric disorders, major medical comorbidity or substance abuse that could interfere with cognitive function, pregnancy and contraindication for MR imaging. Healthy subjects were enrolled after screening for exclusion criteria between July 2012 and January 2016. All subjects underwent a neurological examination and were tested for cognition using the DemTect test (Kalbe et al., 2004) prior to the MR scan. A total of 92 healthy subjects were scanned on a 3 T scanner (Skyra, Siemens, Erlangen) with a 32-channel head coil. The acquisition consisted of 3D T1-weighted MPRAGE (TI=900 ms, α=8°, TE=2.32 ms, TR=2300 ms, GRAPPA=2), 3D T2-weighted SPACE (α=120°, TE=4.08 ms, TR=3200 ms, GRAPPA=2) and 3D inversion recovery prepared SPACE with a FLAIR contrast (TI=1800 ms, α=120°,TE=3. 87 ms, TR=5000 ms, GRAPPA=2) scans with an isotropic resolution of 0.9 mm^3. All data were screened by one of two neuroradiologists (TL, BB) to exclude pathological conditions such as tumor, infarction or other focal lesions. Microangiopathy was not considered a pathological condition, if considered as typical within the age range. Datasets with incomplete data or movement artifacts on visual inspection were also excluded.

Patients

Patients with focal epilepsy undergoing pre-surgical diagnostics were scanned at the same 3T scanner with the same 32-channel head coil and the above mentioned sequences, which is the current standard procedure in our institution. From these patients we retrospectively identified 5 patients with temporal lobe epilepsy and left hippocampal sclerosis (TLE-LHS) and 7 patients with a malformation of cortical development (MCD). The patient data was anonymized before further postprocessing in accordance with the local data protection act ("Landesdatenschutzgesetz"). This approach was approved by the local Ethics committee as post-hoc data analysis.

MRI pre-processing

Both subject and patient scans were processed identically. Initially, the DICOM scans were converted to NIfTI format using MRIConvert (Lewis Center, 2017). VBM analysis was performed using Statistical Parametric Mapping (SPM12) (Wellcome Trust Centre, 2017) running on Matlab R2014b (The MathWorks, Natick, MA). For single-channel processing (classical VBM approach), the converted NIfTI images of the T1 MPRAGE were segmented through unified segmentation using new segment from SPM12 with default settings (bias regularization 0.001, bias cutoff FWHM 60 mm), which generated three main tissue maps in the native space: gray matter (GM_N), white matter (WM_N) and cerebro-spinal fluid (CSF_N) for each subject. The new segment is an extension of the unified segmentation with an ability to use multimodal data for brain tissue segmentation. Multimodal segmentation uses the combination of multiple image contrasts to generate single GM, WM and CSF tissue maps for each subject.

There were 4 possible dual- and triple-modal combinations with the acquired scans and all were evaluated, namely T1+T2, T1+FLAIR, T2+FLAIR and T1+T2+FLAIR. For all the models, T2 and FLAIR images were first co-registered to its corresponding T1 images, using normalized mutual information as the objective function. In the case of the T2+FLAIR model, FLAIR images were co-registered to its corresponding T2 images. To avoid systematic partial volume effects, co-registered images were not resliced prior to segmentation. For every subject, the co-registered image(s) along with the source image T1 (T2 in the case of T2+FLAIR) were simultaneously used as the multi-input channels for unified segmentation generating GM_N, WM_N and CSF_N tissue maps in native space. In both the single- and multi-channel models, a brain template was created in the MNI space with an isotropic resolution of 1.5 mm^3 from the control cohort using DARTEL toolbox. GM_N, WM_N and CSF_N were spatially normalized to the template with DARTEL normalization, resulting in three normalized maps for each subject and patient: GM_c, WM_c, CSF_c that reflects the concentration of the tissue class in any given voxel. As a further step, images were modulated to preserve volume of the native tissue maps within each voxel, resulting in three normalized maps for each subject and patient, GM_v, WM_v, and CSF_v, that reflect the volume of the tissue class in any given voxel. For all scanning sequences, the native structural scans were also normalized to form average T1-weighted, T2-weighted and T2-weighted FLAIR structural templates.

Total intracranial volume

To test the effect of multi-channel segmentation on total intracranial volume (TIV) the volume of the three tissue maps in native space GM_N, WM_N and CSF_N was calculated by the sum over all non-zero voxels multiplied by the voxel volume (0.729 mm^3) resulting in a single volume estimate in ml. TIV was defined as the sum of all three tissue classes. One-way repeated measures ANOVA was used to assess significant differences between the segmentation methods. Mauchly's test was used to test the assumption of sphericity. If the assumption of sphericity was violated, either Greenhouse-Geisser correction (for $\varepsilon \leq .75$) or Huynh-Feldt correction (for $\varepsilon > .75$) was applied. Post-hoc analysis was done with the Bonferroni method.

Table 1
Results of the post-hoc analysis of the repeated measurement ANOVA for in-between group differences for total intracranial volume (TIV) for the different segmentation methods.

TIV		mean difference ± SE calculated against				
Segm. method	Mean (ml) ± SE	T1	T1+T2	T1+ FLAIR	T2+ FLAIR	T1+T2+ FLAIR
T1	1519.8 ± 16.8		82.9 ± 4.2	83.7 ± 3.1	81.3 ± 4.3	78.6 ± 4.2
T1+T2	1436.9 ± 15.4	-82.9 ± 4.2		0.8 ± 2.4	-1.6 ± 0.6	-4.3 ± 0.6
T1+FLAIR	1436.1 ± 16.1	-83.7 ± 3.1	-0.8 ± 2.4		-2.4 ± 2.6	-5.1 ± 2.5
T2+FLAIR	1438.5 ± 15.3	-81.3 ± 4.3	1.6 ± 0.6	2.4 ± 2.6		-2.7 ± 0.4
T1+T2+FLAIR	1441.2 ± 15.4	-78.6 ± 4.2	4.3 ± 0.6	-5.1 ± 2.5	-2.7 ± 0.4	

SE=standard error; Significant results shaded grey, $p < 0.001$.

Differences of the multi-channel segmentation models compared to single-channel segmentation

For the evaluation of the different segmentation models a two-step approach was chosen. In the first step, systematic differences in a paired comparison were identified. For this step the warped modulated subject data GM_v, WM_v and CSF_v were evaluated after smoothing with a 4mm FWHM cubic Gaussian kernel. We compared each multimodal model against T1-only model to account for the GM, WM and CSF changes estimated by all models. For this purpose, general linear models (GLM) were set up for the group concentration and volume analysis of GM, WM and CSF using a paired t-test with a significance threshold of $p < 0.05$ corrected for multiple comparisons with family-wise error rate (FWE). An absolute probability threshold of 0.1 was applied for the exclusion of remote non-gray matter areas. In the second step, regions of significant difference were then evaluated in the individual results and source images of each subject by two neuroradiologists with an expertise in neuroanatomy on MRI scans of 8 years and 10 years (BB, TL) regarding the underlying true anatomy and the resulting segmentation in a consensus approach. This expert visual assessment was based on all three modalities used simultaneously to avoid visual biases.

Age effects

The global volumes of GM, WM and CSF of the different segmentation models were evaluated for a correlation with age after division by TIV to account for different head and brain sizes. Previous studies suggest both a linear and a quadratic component for the change with age for one or more of the tissue classes GM, WM and CSF (Ge et al., 2002; Peelle et al., 2012). Therefore the effect of age on total GM, WM and CSF volume normalized for TIV was analyzed by both a linear model and a second-degree polynomial expansion, to account for both linear and quadratic components of volume change with age. For visualization the second order polynomial fit was compared to the moving average and moving standard deviation (± 1 and 2 SD) with a time span of 5 years.

Identification of hippocampal sclerosis and malformations of cortical development

Local differences between individual patients with TLE-LHS or MCD and the control group were calculated for the unmodulated and modulated data after smoothing with an 8mm FWHM cubic Gaussian kernel. For the TLE-LHS an additional group analysis of all patients vs. controls was calculated. For MCD patients a whole brain analysis was conducted, for patients with TLE-LHS the evaluation was restricted to the anatomical mask of the left hippocampus (taken from the Harvard-Oxford subcortical atlas, provided with FSL; FMRIB, 2017) also smoothed with an 8mm FWHM cubic Gaussian kernel and re-binarized at a threshold of 0.1. General linear models were set up for the in-between group difference of concentration (unmodulated) and volume (modulated) of gray matter, white matter and CSF using an unpaired t-test adjusted for age and gender for each segmentation method. One-way repeated measures ANOVA was used to calculate in-between segmentation methods difference for the T-scores normalized to the individual T-score of the T1-only segmentation. As Mauchly's test indicated that the assumption of sphericity was violated in each case, Greenhouse-Geisser correction (ε was each case $\leq .75$) was applied. Post-hoc analysis was done with the Bonferroni method.

Results

From the 92 datasets of healthy controls, 4 had to be excluded due to incomplete data acquisition for technical reasons (e.g. SAR limitations of the 3D FLAIR), and 11 datasets had to be excluded due to motion artifacts on visual inspection in at least one of the three 3D sequences. All subjects showed a normal clinical examination and a normal cognitive performance in the DemTect. No scores were suspicious for mild cognitive impairment or dementia. A total of 77 datasets, with a complete scan and no movement artifacts on visual inspection could be included in the evaluation, with a mean age of 35.8 years (range of 19–66, 45 female, 32 male).

The 5 patients with TLE-LHS had a mean age of 49.0 (range 25–67, 3 female, 2 male), the 7 patients with MCD had a mean age of 28.1

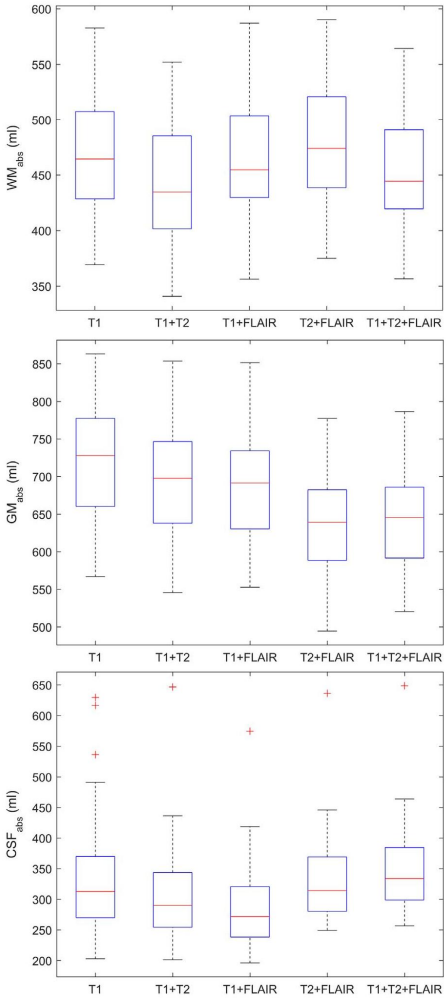


Fig. 2. Boxplots of absolute WM (top), GM (middle) and CSF (bottom) volumes for the different segmentation methods.

(range 19–40, 3 female, 4 male). Of the 7 patients with MCD, one patient with FCD type IIb and one patient with polymicrogyria were pathologically proven. 4 patients had a radiological FCD and one patient had a radiological polymicrogyria.

Absolute and relative volume differences for the different tissue classes and TIV

The mean total intracranial volume over all subjects differed significantly between the segmentation methods (see Fig. 1), $F(1.41, 106.93)=319.63$, $p < 0.001$. The results from the post-hoc analysis showed not only a significant difference between T1-only segmentation and other segmentation methods but also some significant in-between group difference for the other groups (see Table 1).

The absolute volume for WM, GM and CSF are shown in Fig. 2. Repeated measurement ANOVA showed a significant difference between segmentation methods for all three compartments (WM: $F(2.85, 216.72) = 273.96$, $p < 0.001$; GM: $F(2.52, 191.14) = 593.47$, $p < 0.001$; CSF: $F(1.38, 104.76) = 83.29$, $p < 0.001$). Results of the post-hoc comparison are shown in Table 2.

In Figs. 3–5 the results of the paired group comparison of modulated single channel T1 segmentation vs. the different modulated multimodal segmentation methods are illustrated for WM, GM and CSF. There are several regions with significant differences, namely the dura and meninges, the venous sinuses, the occipital lobe, the basal ganglia/thalamus, the pre- and postcentral gyrus, the cerebellar hemispheres, and the pons/medulla oblongata. In a detailed visual analysis of the individual cases by the two neuroradiologists (two typical examples for a 66 year old, and a young 19 year old subject are shown in the Inline Supplemental Figures 1–6) in comparison with the group results, a global effect of single versus multimodal segmentation could be seen. Compared to all other segmentation methods, the T1-only segmentation overestimated GM and CSF outside the brain by falsely applying the GM and CSF tissue class to vessels, meninges, and parts of the dura. This was most pronounced over both hemispheres but also fronto-basal and in the prepontine cistern. Also, the parts of the sagittal, transverse and sigmoid sinus were often treated as CSF and/or GM in the T1-only segmentation. This effect was responsible for the largest part of the approx. 80 ml difference in TIV for T1 vs. the multimodal segmentation results. T1+FLAIR segmentation falsely classified parts of the straight sinus and of the tectum as CSF, which was classified as CSF and GM in the T1-only segmentation (see Inline Supplemental Figure 3, first column).

Besides this global effect there were several anatomical regions with significant difference between T1-only segmentation and some of the multimodal segmentation results. These regions were the infratentorial space with prepontine cistern, pons, medulla oblongata and cerebellum, the occipital lobe, the central region, the region adjacent to the side ventricles, basal ganglia, and thalamus.

In the prepontine cistern and the midpontine angle, flow artifacts in FLAIR and T2-weighted images were challenges for segmentation. The best result for the correct identification of the CSF space in this region was the combination of T1+FLAIR. Only the basilar artery was wrongly assigned to CSF partially. In all other multimodal segmentations larger parts of the CSF space in the peripontine region were allocated to non-brain/non-CSF tissue classes, due to flow artifacts in T2 with hypointense signal instead of bright CSF signal. Segmentation results in the pons and medulla oblongata differed significantly. None of the segmentation methods was able to correctly identify the complex structures in the midbrain. As soon as the T2-weighted image was added to the segmentation some parts of the brain tissue were identified as CSF. A small rim of GM at the border between the pons and the CSF space was falsely identified in the single channel and T1+T2 segmentation. Due to its inhomogeneous contrast large parts of the medulla oblongata were identified as GM as the T2-weighted image was added. In the cerebellum the resolution was not high enough to clearly separate CSF, GM and WM in the source images. Nevertheless, the separation of GM and WM seemed to be more accurate with multimodal segmentation compared to single channel segmentation. The T2-weighted image clearly helped to identify even small areas of CSF and significantly increased CSF content around the cerebellar hemispheres. Whether this is even an overestimation could not be evaluated with the source images. The dentate is not delineated on T1 images and therefore added to WM. Since the standard tissue priors (tissue probability maps) used for segmentation do not account for the cerebellar nuclei even a multimodal segmentation with suitable representation of the dentate on T2 or even better on FLAIR weighted images does not result in a satisfying segmentation of these nuclei.

In the occipital lobe parts of the cortex are less clearly delineated in the T1-weighted images. In this region the single channel segmentation showed a partially disrupted cortical ribbon. In the multimodal segmentation algorithms the cortical ribbon was not disrupted.

Table 2
Results of the post-hoc analysis of the repeated measurement ANOVA for in-between group differences for gray matter (GM) volume, white matter (WM) volume and cerebrospinal fluid (CSF) volume for the different segmentation methods.

WM		mean difference ± SE calculated against				
Segm. method	Mean (ml) ± SE	T1	T1+T2	T1+ FLAIR	T2+ FLAIR	T1+T2+ FLAIR
T1	466.7 ± 6.0		24.6 ± 1.1	1.3 ± 1.0	-11.2 ± 1.4	11.1 ± 1.1
T1+T2	442.1 ± 6.1	-24.6 ± 1.1		-23.3 ± 0.9	-35.8 ± 1.3	-13.5 ± 0.9
T1+FLAIR	465.4 ± 6.1	-1.3 ±1.0	23.3 ± 0.9		-12.5 ± 1.4	9.8 ± 0.6
T2+FLAIR	477.9 ± 6.3	11.2 ± 1.4	35.8 ± 1.3	12.5 ± 1.4		22.3 ± 1.5
T1+T2+FLAIR	455.6 ± 5.9	-11.1 ± 1.1	13.5 ± 0.9	-9.8 ±0.6	-22.3 ± 1.5	
GM						
T1	722.0 ± 8.6		29.1 ± 2.5	35.5 ± 2.8	87.7 ± 2.5	80.2 ± 2.8
T1+T2	692.9 ± 8.2	-29.1 ± 2.5		6.4 ± 2.0[a]	58.7 ± 1.6	51.1 ± 2.1
T1+FLAIR	686.5 ± 7.5	-35.5 ± 2.8	-6.4 ± 2.0[a]		52.3 ± 1.8	44.7 ± 1.0
T2+FLAIR	634.3 ± 7.6	-87.7 ± 2.5	-58.7 ± 1.6	-52.3 ± 1.8		-7.5 ± 1.5
T1+T2+FLAIR	641.8.6 ± 7.1	-80.2 ± 2.8	-51.1 ± 2.1	-44.7 ± 1.0	7.5 ± 1.5	
CSF						
T1	331.1 ± 9.6		29.2 ± 5.5	47.0 ± 4.8	4.8 ± 5.7	-12.7 ± 5.7
T1+T2	301.9 ± 7.7	-29.2 ± 5.5		17.7 ± 2.4	-24.5 ± 1.3	-41.9 ± 1.4
T1+FLAIR	284.2 ± 7.2	-47.0 ± 4.8	-17.7 ± 2.4		-42.2 ± 2.3	-59.6 ± 2.3
T2+FLAIR	326.4 ± 7.0	-4.8 ± 5.7	24.5 ± 1.3	42.2 ± 2.3		-17.4 ± 0.7
T1+T2+FLAIR	343.8 ± 6.8	12.7 ± 5.7	41.9 ± 1.4	59.6 ± 2.3	17.4 ± 0.7	

SE=standard error; Significant results shaded grey, $p < 0.001$, if not stated otherwise: [a]$p=0.021$.

Compared to the other segmentation methods T2+FLAIR overestimated the thickness of the cortical ribbon in the occipital lobe and parts of the CSF space and the dura were falsely classified as WM. T1+T2 segmentation seemed to show the best anatomical detail of the delicate visual cortex.

At the medial border of the lateral ventricles and at the border of the tentorium a thin rim of GM was identified in the single channel segmentation. In all other segmentations these regions were correctly not identified as GM. At the rim of the corpus callosum the partial volume of CSF was in all multimodal segmentations that included a T2-weighted image (T1+T2, T2+FLAIR, T1+T2+FLAIR) larger than for the single channel segmentation and T1+FLAIR.

Around the central sulcus several changes to the segmentation were noted compared to T1-only segmentation, as soon as a FLAIR image was included. These effects were most pronounced in T2+FLAIR segmentation. WM had a significant larger volume in both the primary

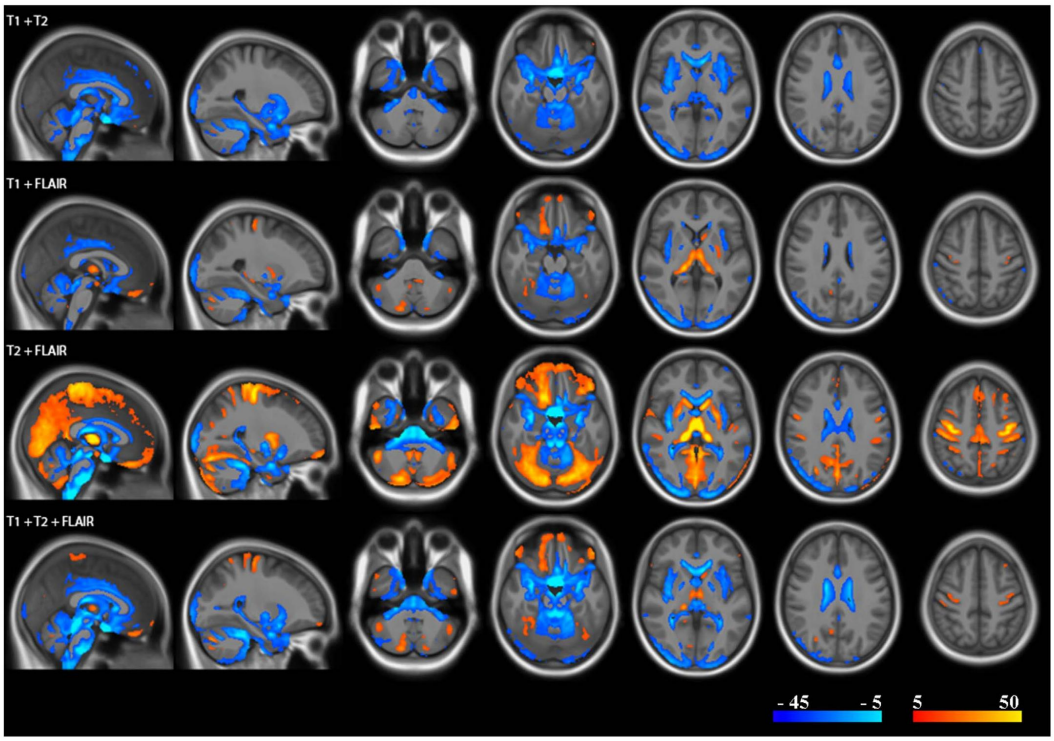


Fig. 3. Results of the paired t-tests of multimodal segmentation models vs. T1-only segmentation calculated on the modulated warped WM images. Each row represents the results of one multimodal segmentation combination (same 7 slices shown in the Inline Supplemental Figure 1–6). In blue areas with significantly less WM and in red-yellow significantly more WM compared to T1-only segmentation are shown. The t-score is coded in the color intensity.

motor cortex (PMC, precentral gyrus) and the primary sensory cortex (PSC, postcentral gyrus) while the GM volume was reduced. Based on the anatomical T1 weighted scan GM seemed underestimated in multimodal segmentations with FLAIR included to both expert raters.

In the basal ganglia none of the segmentation algorithms perfectly delineated GM and WM structures. Single channel segmentation seemed to be superior in most aspects. GM volume was falsely significantly smaller for the globus pallidus and the lateral parts of the thalamus for T1+FLAIR, T2+FLAIR and T1+T2+FLAIR segmentation. The central part of the globus pallidus was even classified as CSF in T1+FLAIR, T2+FLAIR and T1+T2+FLAIR segmentation. Even parts of the thalamus were falsely classified as CSF in T2+FLAIR and T1+T2+FLAIR segmentation, while T1+FLAIR and T2+FLAIR identified larger parts of the medial thalamus as WM instead of GM. Only at the claustrum did WM seem to be slightly overestimated in the single channel segmentation, which was better identified in the multimodal methods. T2+FLAIR segmentation misclassified parts of the caudate nucleus and the putamen as WM instead of GM.

Age effects

There was a strong negative linear relationship between age and relative GM volume for all segmentation methods (see Fig. 6, first column and Table 3). T1+T2+FLAIR segmentation showed the best fit of the data to a linear model, followed by T1+FLAIR and T2+FLAIR (see Table 3). Relative WM showed only a slight volume change with age. After an initial increase in relative WM volume up to the age of approx. 45 years the WM volume showed a slow decrease (see Fig. 6, second column). The statistical evaluation showed both a linear increasing component and a quadratic decreasing component (see Table 2), although significance was slightly missed for the quadratic term for T1+T2 and T1+T2+FLAIR segmentation. All the segmentation methods had a rather similar goodness of fit of the data towards the second order polynomial (see Table 4). Relative CSF volume showed a clear increase with age and a trend towards a faster increase in volume in older subjects (see Fig. 6, third column). There was one outlier with a notably larger relative CSF volume at age 62, which increased the moving standard deviation for the later age range. While the moving average shows a clear deviation from a pure linear trend a second order polynomial did not better explain the data and failed to reach significance (see Table 5). T1+T2+FLAIR segmentation showed the best fit of the data for the linear and second order polynomial model, with T1+FLAIR and T2+FLAIR performing almost as well (see Table 5).

Identification of focal lesions (MCD and TLE-LHS)

Table 6 shows the results of the T-scores in the individual lesion normalized by the T-score of the T1-only segmentation. For the patients with MCD the FLAIR image had the strongest influence on the T-score with either T1+FLAIR, T2+FLAIR or T1+T2+FLAIR segmentation resulting in the highest T-score, although a one-way repeated-measure ANOVA showed no significant differences between segmentation methods for the individual T-scores.

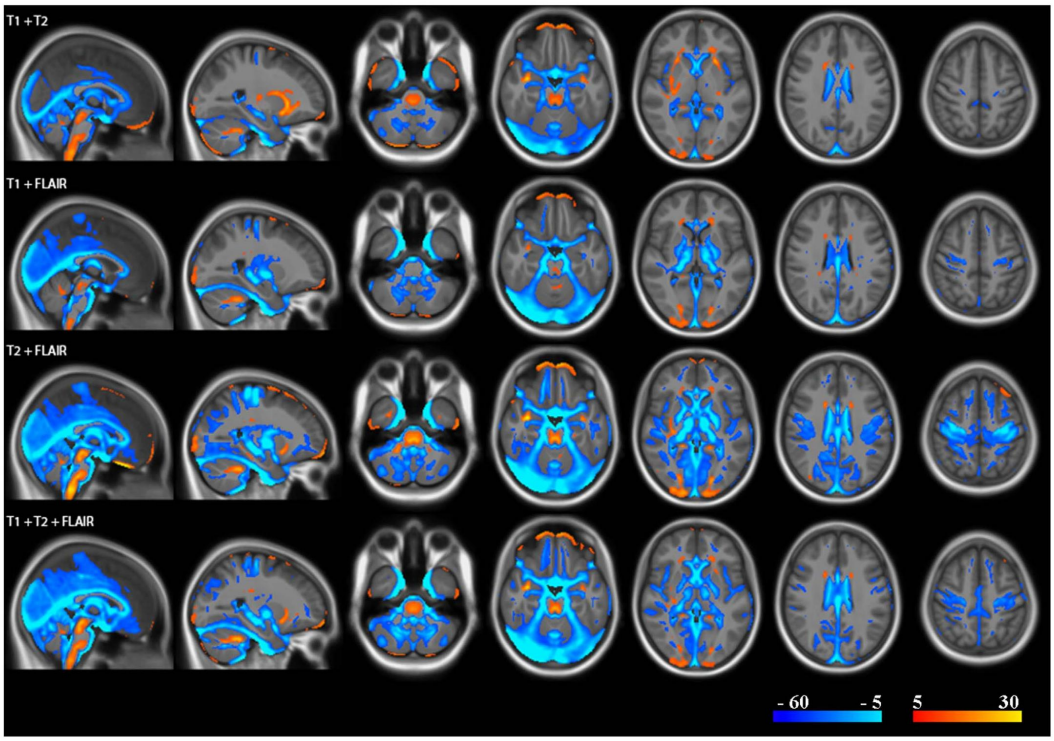


Fig. 4. Results of the paired t-tests of multimodal segmentation models vs. T1-only segmentation calculated on the modulated warped GM images. Each row represents the results of one multimodal segmentation combination (same 7 slices shown in the Inline Supplemental Figure 1–6). In blue areas with significantly less GM and in red-yellow significantly more GM compared to T1-only segmentation are shown. The t-score is coded in the color intensity.

In contrast, in patients with TLE-LHS the combination of T1+T2 was clearly superior in comparison to the other segmentation methods in the individual cases and in the group analysis (sees Table 7).

Discussion

This is the first study on the effect of multimodal segmentation in SPM12 based on 3D- T1, T2 and FLAIR images in a larger cohort to our knowledge. There was a clear and significant difference in the segmentation results of the five tested segmentation methods. Total intracranial volume showed the highest values for the single channel segmentation based on T1-weighted scans by around 80 ml (~5%) on average due to misclassification of dura matter and vessels to GM and to CSF, in line with previous reports (van der Kouwe et al., 2008; Viviani et al., 2016). A recent study compared SPM12 segmentation, SPM8 segmentation and FreeSurfer segmentation with the gold standard manual delineation for the assessment of TIV (Malone et al., 2015). They found the best performance for SPM12, which had a slight underestimation of the "real" TIV defined by manual delineation by around 40 ml. They used a slightly different approach to define TIV, by inverse-transforming a template-space TIV mask to the individual (Malone et al., 2015), but a previous study showed no differences between this approach and the combination of the modulated warped GM, WM and CSF maps for SPM8 (Ridgway et al., 2011). As TIV was previously defined as all structures within the skull including meninges and vessels (e.g. (Malone et al., 2015; Whitwell et al., 2001)), TIV estimation by multi-channel segmentation and combination of the modulated warped GM, WM and CSF maps probably underestimates the "real" TIV as it better removes vessels and meninges from the tissue classes GM, WM and CSF. Vessels and meninges only occupy a small volume of the TIV and their volume should be not affected by atrophy in neurodegenerative disorders or most other neuro-psychiatric conditions. A small underestimation of the "real" TIV may, in theory, lead to a slightly increased sensitivity to atrophy effects by removing a potentially unrelated factor (total volume of vessels and meninges) from the analysis. However, it is not known whether this has practical consequences. For longitudinal studies, Whitwell et al. suggest using the same acquisition for defining TIV as that used for segmentation (Whitwell et al., 2001). More worrisome is the misclassification of vessels and meninges as GM in T1-only segmentation. If one is interested in GM atrophy a misclassification of vessels and meninges should decrease the sensitivity for the identification of nearby atrophy and there may be false positive results if one is interested in MCD with an expected focal increase of GM. This misclassification of GM might also explain the larger standard deviation in elderly subjects in the fit of age vs. GM volume fraction for T1-only segmentation (Fig. 6).

As the data of the different segmentation models had the same unit and were fitted to the same dependent variable age one can directly compare the standard error of the regression (root-mean-squared error). Compared to adjusted R^2 it is the better predictor of the goodness of fit, as R^2 only measures the percentage reduction in mean-squared-error of the regression model relative to a model that is defined by a constant term. While for WM no real difference in the

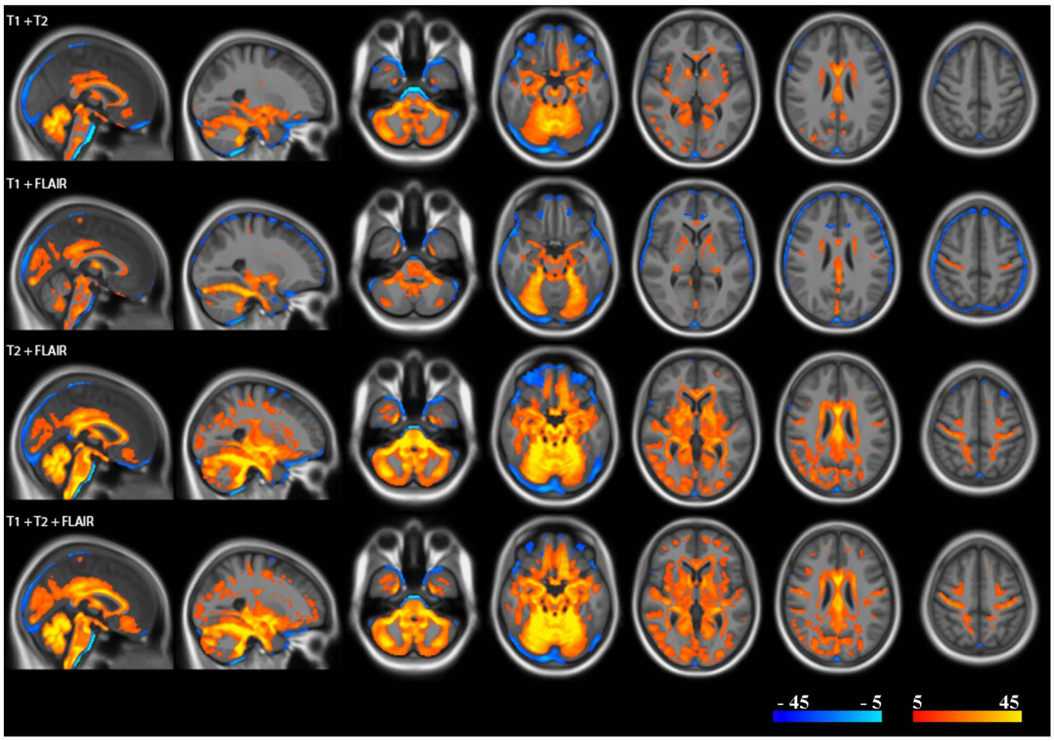


Fig. 5. Results of the paired t-tests of multimodal segmentation models vs. T1-only segmentation calculated on the modulated warped CSF images. Each row represents the results of one multimodal segmentation combination (same 7 slices shown in the Inline Supplemental Figure 1–6). In blue areas with significantly less CSF and in red-yellow significantly more CSF compared to T1-only segmentation are shown. The t-score is coded in the color intensity.

standard error of the regression could be found, the combination of T1+ T2 + FLAIR had the best fit for GM and CSF followed by T1+FLAIR which only performed slightly worse. This is in good agreement with the visual rating by the two neuroradiologists, who judged T1+T2+FLAIR and T1+FLAIR to best represent the anatomy for the supratentorial brain, excluding the basal ganglia. The basal ganglia showed the best segmentation results in the single channel T1-segmentation. Decreasing signal intensity of the basal ganglia on both T2- and FLAIR-weighted scans, namely the globus pallidum and the pulvinar, with increasing age, led to a significant misclassification of these structures. But even the single-channel T1-segmentation showed clear segmentation errors, which is a well-known problem due to the low contrast in these iron rich structures (Lorio et al., 2016). Lorio et al. recently suggested using new tissue priors for subcortical structures to improve segmentation in this region (Lorio et al., 2016). The performance of these new priors in relation to multi-spectral segmentation needs to be evaluated in the future. Although, T1+FLAIR and T1+T2+FLAIR had similar results for the correlation with age, there was a clear difference in the absolute volume fractions (see Fig. 2). Adding the T2-weighted scan to T1+FLAIR increased the CSF volume fraction at the cost of the GM volume. However, this difference was so small that it was not evident on visual inspection and was only detectable in group level comparison.

Significant differences were seen in the primary motor cortex, the primary sensory cortex and the primary visual cortex for the different segmentation models. In these areas the myelin content of GM is larger, which can be made visible by dividing a T1-weighted image by a T2-weighted image (Glasser, Van Essen, 2011). Nevertheless, the effect was different for the primary visual cortex and for the PMC and PSC. Around the central sulcus WM volume increased as soon as a FLAIR image was added, while in the occipital lobe GM increased as soon as either T2 or FLAIR was added to the segmentation. This effect could be explained by the very thin cortical ribbon of the occipital lobe (Glasser et al., 2016), and its rather low GM-WM-contrast which is well known from routine imaging.

All models failed to segment the brainstem and the surrounding CSF correctly. A previous study yielded good results by adding a PD-weighted scan for segmentation of the brainstem (Lambert et al., 2013), which was not available in our current dataset. Neither FLAIR nor T2-weighted images seem an adequate replacement of PD-weighted scans in the brainstem. In the cerebellum the addition of a second contrast significantly increased the CSF fraction; this was especially true for the T2-weighted scan but to a lesser extent also for the FLAIR-weighted scan. As the cerebellar cortex has a width of around 0.4 to 0.8 mm (Haque, 2012; Marques et al., 2010) every voxel of the sequences used is always a mixture of GM with CSF and/or WM, and the “truth” cannot be easily determined visually. To assess the best segmentation results for the cerebellum an ultra-high resolution reference scan would be needed for reference. With the increasing availability of ultra-high field scanners this might be available in the near future.

In good agreement with a variety of previous cross-sectional and longitudinal studies (Ge et al., 2002; Good et al., 2001; Hedman et al., 2012; Kruggel, 2006; Peelle et al., 2012; Riello et al., 2005) age

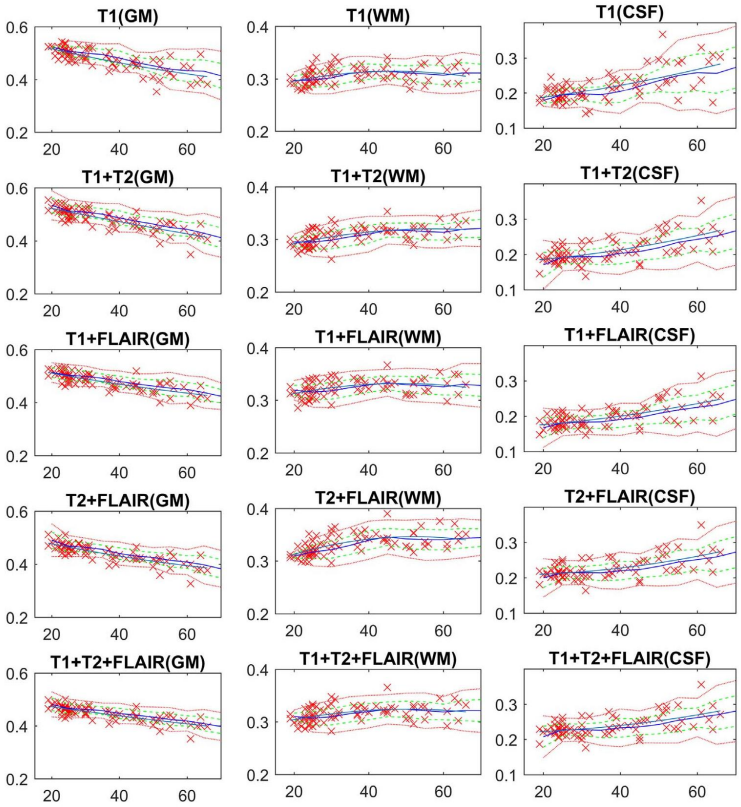


Fig. 6. Scatter plot of age (x-axis) vs. relative volume fraction (y-axis) for GM (first column), WM (second column) and CSF (third column) for T1-only segmentation (first row), and multimodal segmentation models: T1+T2 (second row), T1+FLAIR (third row), T2+FLAIR (fourth row), T1+T2+FLAIR (fifth row). The blue line represents the moving average with a time span of 5 years, with 1 standard deviation (dashed green line) and 2 standard deviations (dotted red line). The turquoise line represents the 2nd order polynomial fit.

Table 3
Fitting results of a second order polynomial model (LP = linear and QP = quadratic part) and a linear model between age and relative tissue volume for GM.

Segm. method	Adj. R^2	Root-mean-squared error	Beta (LP)	p-value (LP)	Beta (QP)	p-value (QP)
T1	0.578 0.582	0.027 0.027	−0.00346 −0.00235	**0.038** **<** **0.001**	0.00001	0.499
T1+T2	0.649 0.653	0.023 0.023	−0.00272 −0.00238	0.061 **<** **0.001**	0.00000	0.814
T1+FLAIR	0.599 0.602	0.020 0.020	−0.00268 −0.00187	**0.035** **<** **0.001**	0.00001	0.516
T2+FLAIR	0.637 0.639	0.021 0.021	−0.00308 −0.00204	**0.017** **<** **0.001**	0.00001	0.411
T1+T2+FLAIR	0.594 0.598	0.018 0.018	−0.00203 −0.00162	0.067 **<** **0.001**	0.00001	0.706

Table 4
Fitting results of a second order polynomial model (LP=linear and QP=quadratic part) and a linear model between age and relative tissue volume for WM.

Segm. method	Adj. R^2	Root-mean-squared error	Beta (LP)	p-value (LP)	Beta (QP)	p-value (QP)
T1	0.114 0.063	0.015 0.016	0.00244 0.00033	**0.010** **0.016**	−0.00003	**0.024**
T1+T2	0.284 0.261	0.015 0.015	0.00233 0.00068	**0.012** **<** **0.001**	−0.00002	0.069
T1+FLAIR	0.127 0.087	0.015 0.015	0.00230 0.00038	**0.014** **0.005**	−0.00002	**0.038**
T2+FLAIR	0.321 0.262	0.016 0.017	0.00342 0.00076	**0.001** **<** **0.001**	−0.00003	**0.008**
T1+T2+FLAIR	0.148 0.117	0.015 0.015	0.00216 0.00043	**0.019** **0.001**	−0.00002	0.057

Table 5
Fitting results of a second order polynomial model (LP=linear and QP=quadratic part) and a linear model between age and relative tissue volume for CSF.

Segm. method	Adj. R^2	Root-mean-squared error	Beta (LP)	p-value (LP)	Beta (QP)	p-value (QP)
T1	0.394 0.400	0.033 0.033	0.00101 0.00203	0.619 < **0.001**	0.00001	0.616
T1+T2	0.405 0.408	0.027 0.027	0.00039 0.00171	0.816 < **0.001**	0.00002	0.431
T1+FLAIR	0.393 0.397	0.025 0.025	0.00038 0.00150	0.802 < **0.001**	0.00001	0.455
T2+FLAIR	0.324 0.323	0.025 0.025	−0.00034 0.00129	0.820 < **0.001**	0.00002	0.278
T1+T2+FLAIR	0.299 0.301	0.024 0.024	−0.00013 0.00119	0.930 < **0.001**	0.00002	0.368

Table 7
VBM findings for the group level analysis of gray matter concentration and volume in patients with LHS in the left hippocampus region with small volume correction.

Modality	**Cluster size p < 0.05 (FWE)**	**Cluster size p < 0.0001 (uncorr.)**	**Cluster size p < 0.001 (uncorr.)**	**Tmax (T value map)**
T1 concentration	0	0	151	3.83
T1+T2 concentration	41	370	773	5.94
T1+FLAIR concentration	0	0	12	3.46
T2+FLAIR concentration	0	14	161	4.11
T1+T2+FLAIR concentration	0	0	35	3.83
T1 volume	0	0	122	3.69
T1+T2 volume	0	242	873	4.8
T1+FLAIR volume	0	0	1	3.42
T2+FLAIR volume	0	11	272	4.02
T1+T2+FLAIR volume	0	0	39	3.45

dependent atrophy in healthy adults is mainly driven by a decrease of GM in our sample. Our data suggests a linear decrease in the age range examined, which is also in good agreement with previous data. Previous examinations suggest an increase of WM volume up to the age of around 45 (Hedman et al., 2012; Kruggel, 2006; Peelle et al., 2012) and then a steady decrease. This U-shaped curve is also present in our data, although significance is missed probably due to the low number of subjects with an age > 45 years. The combination of the volume change effects of GM and WM leads to a relatively flat curve of the CSF volume fraction vs. age up to the age of around 30 and then a steady increase of the CSF volume fraction with a larger standard deviation of the residuals with increasing age. This suggests that a pure 2^{nd} order polynomial or linear model is not a good descriptor of the data. Compared to all multimodal segmentation methods the correlation of age with GM showed an increased standard deviation with age. A limitation of this study is the missing gold standard. Therefore the "true" GM, WM and CSF volume is not known. However, the sample consisted of healthy adults with no known or symptomatic neurological disorders and were tested for MCI. This suggests that the standard deviation for a random sample should have a normal distribution and should not change with increasing age. Also, the misclassification of vessels and dura mater in T1-only segmentation is well appreciated on the scans and is a good explanation for this effect.

A first test on individual patient data showed a clear trend that favored a multimodal segmentation approach with increasing T-values for both MCD and TLE-LHS. Depending on the individual pathology and, probably, its contrast on conventional images, the best combination of image contrast differs. For MCD the addition of a 3D FLAIR to a T1 MPRAGE seems to be the most important step, while in hippocampal sclerosis a T2-weighted 3D SPACE seems superior.

The increase in sensitivity of multimodal segmentation comes at the cost of longer measurement time, as each 3D sequence needs about 5–7 minutes' acquisition time. If one has the time to invest an additional 5–7 minutes the addition of a T2-weighted or FLAIR-weighted 3D

Table 6
Multispectral VBM findings of gray matter concentration (no modulation) and volume (modulation applied) for individual patients compared against the control cohort. In the case of LHS, all the data is small volume corrected after the application of an inclusive mask of left hippocampus. T-scores were normalized by the Tmax of T1-only segmentation, raw Tmax for T1-only segmentation given in parentheses. The highest overall T-scores are shown in bold.

	Concentration					**Volume**				
	T1	**T1 + T2**	**T1 +FLAIR**	**T2 +FLAIR**	**T1 + T2 +FLAIR**	**T1**	**T1 + T2**	**T1 +FLAIR**	**T2 +FLAIR**	**T1 + T2 +FLAIR**
MCD (PM)	1 (19.12)	1.02	1.14	**1.24**	1.18	1 (4.88)	1.04	1.17	1.24	**1.25**
MCD (FCD)	1 (15.86)	0.98	1.12	**1.23**	1.03	1 (4.87)	1.08	**1.14**	1.13	**1.14**
MCD (FCD)	1 (17.65)	1.03	**1.19**	1.16	1.14	1 (4.91)	0.91	**1.11**	1.02	1.08
MCD (FCD)	1 (20.43)	0.88	0.97	**1.13**	0.83	1 (4.14)	1.00	0.99	**1.27**	1.03
MCD (FCD)	1 (2.27)	0.96	**1.33**	1.14	1.25	1 (1.17)	1.14	1.48	**1.66**	1.56
MCD (PM)	1 (18.94)	1.09	1.28	**1.30**	1.27	1 (10.56)	1.29	1.29	1.26	**1.45**
MCD (FCD)	1 (3.25)	0.92	**1.40**	0.94	1.30	1 (4.25)	1.01	1.02	1.00	**1.11**
LHS	1 (2.21)	1.06	1.02	**1.16**	0.97	**1 (2.18)**	0.71	0.66	0.79	0.54
LHS	1 (7.70)	**1.11**	0.74	0.81	0.71	1 (3.93)	**1.13**	0.95	1.06	1.00
LHS	1 (4.83)	**1.17**	0.76	1.16	0.74	**1 (5.22)**	0.87	0.69	0.91	0.70
LHS	1 (2.01)	**1.46**	1.31	1.17	1.09	1 (1.60)	**1.53**	1.11	1.16	1.52
LHS	1 (5.28)	**2.47**	1.01	2.02	1.75	1 (3.00)	**1.61**	1.03	1.47	1.36

MCD – malformation of cortical development; LHS – left hippocampal sclerosis; PM – polymicrogyria; FCD – focal cortical dysplasia

sequence could be beneficial depending on the clinical or scientific question/interest. As movement artifacts also have a large influence, it needs to be evaluated if a double acquisition of a T1-weighted sequence alone could reduce the number of drop outs and could therefore be more beneficial depending on the patient cohort (e.g. children or patients with dementia). Furthermore, small misalignments between different sequences (e.g. residual movement, differences in image distortions) cannot be ruled out and might impact on multimodal compared to single channel segmentations in certain brain areas. Moreover, the "ground truth", in our case expert visual assessment, is naturally limited to few sequences. However, in our study all three modalities were used simultaneously to avoid visual biases.

As a further option, different machine learning algorithms have been recently proposed for brain segmentation (e.g. LINKS (Wang et al., 2015) or PyraMiD-LSTM (Stollenga et al., 2015)). Compared to SPM12 multi-channel segmentation based on T1 and FLAIR images, many of these algorithms perform better (see MRBrainS, 2017 for the current results of different algorithms) on a dataset that consists of 5 training datasets and 15 test datasets with 3D T1 (1mm³ isotropic resolution), 2D FLAIR (0.958×0.958×3mm³ resolution) and 2D IR-T1 (0.958×0.958×3mm³ resolution) of elderly subjects (age > 50 years), which is publicly available (MRBrainS, 2017). It is possible that such algorithms may also perform better in our cohort. However, the main disadvantage is that machine learning approaches usually need a training dataset e.g. of manually segmented images. To create high quality manual segmented datasets is very time consuming especially on high resolution images and with multiple contrasts. Furthermore, if different contrasts are used, the manual segmentations of each contrast might not converge, as manual segmentation is usually binary task in which every voxel can only be assigned to one tissue class. If manual segmentation is only based on one contrast, the training datasets will favor this contrast in a multi-channel segmentation. Therefore, it is difficult to identify the algorithm that truly segments best. In some regions, like the thalamus or the basal ganglia, and in identifying microangiopathy machine learning algorithms have an advantage over methods that just consider pre-defined tissue classes (GM, WM and CSF) like standard SPM12 analysis. Further studies could objectively compare the performance of the frequently used SPM-based segmentation with such advanced models, preferably integrated into publicly available initiatives like MRBrainS.

In conclusion, a multimodal segmentation with T1+T2+FLAIR or T1+FLAIR contrast showed the best overall segmentation results for gray matter, white matter and CSF. For cortical GM segmentation, if only two sequences can be used the combination of T1+FLAIR seems to be superior to T1+T2 or T2+FLAIR in general, but filigree structures like the occipital visual cortex or the hippocampus are best delineated on T1+T2. For CSF, a combination of T1+FLAIR seems to be the best. Single channel T1 segmentation classifies parts of dura and vessels as GM and CSF, which leads to a significant increase of TIV by about 80 ml. This misclassification could decrease significance in atrophy detection, might cause false positive findings in lesion detection in MCD, and could explain the larger standard deviation for the correlation with age in elderly subjects.

For basal ganglia and infratentorial segmentation none of the methods yielded a sufficient result. Recently published new tissue priors (Lorio et al., 2016) for improved segmentation of subcortical structures like the basal ganglia, the thalamus or the dentate may resolve this shortcoming. Further evaluation with these tissue priors in a multimodal approach is therefore needed. The effect of changing the standard parameters, e.g. for bias regularization and bias cutoff were not tested in our study. As the default settings are recommended in SPM12 and used in the vast majority of SPM-based studies, our results should be generalizable to most VBM analyses done with 1.5 or 3 T and standard full-brain coils. With "atypical" scenarios like high-field MRI (7 T and above) or partial coverage head coils, different bias settings may be needed that could impact also on multimodal segmentation. Furthermore, the effect of additional variance from different scanners is well known for T1-weighted VBM approaches (Focke et al., 2011; Takao et al., 2013). Whether a multi-channel segmentation also reduces the effect of scanner variability or even multiplies this effect needs to be evaluated in further studies.

For the identification of pathologic lesions, in our case two typical lesions in focal epilepsy, different combinations yield the best results, probably depending on the type of lesion, the underlying histopathology and, potentially, the anatomical region. In hippocampal sclerosis an additional T2-weighted scan is useful, whereas for the detection of cortical malformations T2-FLAIR is the better option. It is likely that similar differentiations can be elucidated for other pathological conditions.

In summary, multi-modal models can improve the segmentation quality compared to a T1-only approach. However, the exact choice of sequence combinations has to be tailored to the area of interest and the expected pathology.

Funding

This study received research funding by the Medical Faculty of the University of Tübingen (grant number AKF 321-0-1).

Additional information

The T1, T2 and FLAIR datasets of the normal controls will be made available by the authors on request. Requests can be made via Email to either Benjamin Bender (Benjamin.Bender@med.uni-tuebingen.de), Tobias Lindig (Tobias.Lindig@med.uni-tuebingen.de) or Niels Focke (Niels.Focke@uni-tuebingen.de).

Acknowledgment

The authors wish to thank Merim Bilalic and Matthew Bladen for proof-reading of the manuscript and language editing.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.neuroimage.2017.02.016.

References

Ashburner, J., Friston, K.J., 2005. Unified segmentation. Neuroimage 26, 839–851. http://dx.doi.org/10.1016/j.neuroimage.2005.02.018.

Diaz-de-Grenu, L.Z., Acosta-Cabronero, J., Pereira, J.M., Pengas, G., Williams, G.B., Nestor, P.J., 2011. MRI detection of tissue pathology beyond atrophy in Alzheimer's disease: introducing T2-VBM. Neuroimage 56, 1946–1953. http://dx.doi.org/10.1016/j.neuroimage.2011.03.082.

Focke, N.K., Helms, G., Kaspar, S., Diederich, C., Toth, V., Dechent, P., Mohr, A., Paulus, W., 2011. Multi-site voxel-based morphometry–not quite there yet. Neuroimage 56, 1164–1170. http://dx.doi.org/10.1016/j.neuroimage.2011.02.029.

Ge, Y., Grossman, R.I., Babb, J.S., Rabin, M.L., Mannon, L.J., Kolson, D.L., 2002. Age-related total gray matter and white matter changes in normal adult brain. Part I: volumetric MR imaging analysis. AJNR Am. J. Neuroradiol. 23, 1327–1333.

Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., Smith, S.M., Van Essen, D.C., 2016. A multi-modal parcellation of human cerebral cortex. Nature 536, 171–178. http://dx.doi.org/10.1038/nature18933.

Glasser, M.F., Van Essen, D.C., 2011. Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. J Neurosci. 31, 11597–11616. http://dx.doi.org/10.1523/jneurosci.2180-11.2011.

Good, C.D., Johnsrude, I.S., Ashburner, J., Henson, R.N., Friston, K.J., Frackowiak, R.S., 2001. A voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage 14, 21–36. http://dx.doi.org/10.1006/nimg.2001.0786.

Haque, A., 2012. Study on thickness of gray matter of cerebellum in Bangladeshi Cadaver. Bangladesh J. Anat. 10, 32–36.

Hedman, A.M., van Haren, N.E., Schnack, H.G., Kahn, R.S., Hulshoff Pol, H.E., 2012. Human brain changes across the life span: a review of 56 longitudinal magnetic resonance imaging studies. Hum. Brain Mapp. 33, 1987–2002. http://dx.doi.org/10.1002/hbm.21334.

Kalbe, E., Kessler, J., Calabrese, P., Smith, R., Passmore, A.P., Brand, M., Bullock, R.,

2004. DemTect: a new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. Int. J. Geriatr. Psychiatry 19, 136–143. http://dx.doi.org/10.1002/gps.1042.

Kruggel, F., 2006. MRI-based volumetry of head compartments: normative values of healthy adults. Neuroimage 30, 1–11. http://dx.doi.org/10.1016/j.neuroimage.2005.09.063.

Lambert, C., Lutti, A., Helms, G., Frackowiak, R., Ashburner, J., 2013. Multiparametric brainstem segmentation using a modified multivariate mixture of Gaussians. Neuroimage Clin. 2, 684–694. http://dx.doi.org/10.1016/j.nicl.2013.04.017.

Lorio, S., Fresard, S., Adaszewski, S., Kherif, F., Chowdhury, R., Frackowiak, R.S., Ashburner, J., Helms, G., Weiskopf, N., Lutti, A., Draganski, B., 2016. New tissue priors for improved automated classification of subcortical brain structures on MRI. Neuroimage 130, 157–166. http://dx.doi.org/10.1016/j.neuroimage.2016.01.062.

Malone, I.B., Leung, K.K., Clegg, S., Barnes, J., Whitwell, J.L., Ashburner, J., Fox, N.C., Ridgway, G.R., 2015. Accurate automatic estimation of total intracranial volume: a nuisance variable with less nuisance. Neuroimage 104, 366–372. http://dx.doi.org/10.1016/j.neuroimage.2014.09.034.

Marques, J.P., van der Zwaag, W., Granziera, C., Krueger, G., Gruetter, R., 2010. Cerebellar cortical layers: in vivo visualization with structural high-field-strength MR imaging. Radiology 254, 942–948. http://dx.doi.org/10.1148/radiol.09091136.

Peelle, J.E., Cusack, R., Henson, R.N., 2012. Adjusting for global effects in voxel-based morphometry: gray matter decline in normal aging. Neuroimage 60, 1503–1516. http://dx.doi.org/10.1016/j.neuroimage.2011.12.086.

Ridgway, G.R., Barnes, J., Pepple, T., Fox, N.C., 2011. Estimation of Total Intracranial Volume; A Comparison of Methods. 2011 7. AAICAD Alzheimer's Imaging Consortium, Paris, France, S62–S63.

Riello, R., Sabattoli, F., Beltramello, A., Bonetti, M., Bono, G., Falini, A., Magnani, G., Minonzio, G., Piovan, E., Alaimo, G., Ettori, M., Galluzzi, S., Locatelli, E., Noiszewska, M., Testa, C., Frisoni, G.B., 2005. Brain volumes in healthy adults aged 40 years and over: a voxel-based morphometry study. Aging Clin. Exp. Res. 17, 329–336.

Stollenga, M.F., Byeon, W., Liwicki, M., Schmidhuber, J., 2015. . Parallel multi-dimensional LSTM, with application to fast biomedical volumetric image segmentation. In: Proceedings of the 28th International Conference on Neural Information Processing Systems, Montreal, Canada, pp. 2998–3006.

Takao, H., Hayashi, N., Ohtomo, K., 2013. Effects of the use of multiple scanners and of scanner upgrade in longitudinal voxel-based morphometry studies. J. Magn. Reson. Imaging 38, 1283–1291. http://dx.doi.org/10.1002/jmri.24038.

van der Kouwe, A.J., Benner, T., Salat, D.H., Fischl, B., 2008. Brain morphometry with multiecho MPRAGE. Neuroimage 40, 559–569. http://dx.doi.org/10.1016/j.neuroimage.2007.12.025.

Viviani, R., Pracht, E., Brenner, D., Stingl, J., Stöcker, T., 2016. . Multimodal segmentation utilizing FLAIR or R2* images for improved detection of gray matter in VBM. In: Proceedings of the 22nd Annual meeting of the Organization for Human Brain Mapping, Geneva, Switzerland, 2197.

Wang, L., Gao, Y., Shi, F., Li, G., Gilmore, J.H., Lin, W., Shen, D., 2015. LINKS: learning-based multi-source IntegratioN frameworK for Segmentation of infant brain images. Neuroimage 108, 160–172. http://dx.doi.org/10.1016/j.neuroimage.2014.12.042.

Whitwell, J.L., Crum, W.R., Watt, H.C., Fox, N.C., 2001. Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging. AJNR Am. J. Neuroradiol. 22, 1483–1489.

Web references

Lewis Center, 2017. for Neuroimaging, University of Oregon, USA. MRIConvert. ⟨https://lcni.uoregon.edu/downloads/mriconvert/⟩ (last accessed 22.01).

Wellcome Trust Centre, 2017. for Neuroimaging, London, UK. Statistical Parameter Mapping (SPM), ⟨http://www.fil.ion.ucl.ac.uk/spm/⟩ (last accessed 22.01).

FMRIB, O., 2017. , UK. FMRIB Software Library. ⟨https://fsl.fmrib.ox.ac.uk/fsl/⟩ (last accessed 22.01).

MRBrainS, 2017. , University Medical Center Utrecht, Netherlands. Segmentation results of the Grand Challenge on MR Image Segmentation workshop. ⟨http://mrbrains13.isi.uu.nl/results.php⟩ (last accessed 22.01).

4.2. Study II (in press)

Systematic Assessment of Multispectral Voxel-Based Morphometry in MRI Previously Negative for 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 MR imaging previously negative for focal epilepsy.

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 Magentom Skyra 3T scanner with an isotropic resolution of 0.9 mm^3. 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 patients with MR imaging negative for focal epilepsy.

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.[1] A common epileptogenic lesion is focal cortical dysplasia, which has notable morphologic characteristics on MR imaging.[2] However, pure visual analysis, especially in subtle cases, can be challenging. Also, many histopathologically proved focal cortical dysplasias escape visual detection.[3] Failure to identify these lesions can often label patients as having MR imaging negative for focal cortical dysplasia. These patients have poor surgical outcomes or may not even be referred for a potential epilepsy operation.[4]

During the past decade, multiple MR imaging postprocessing methods have been applied to improve lesion detection in epilepsy,[5] 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.[6] This approach is commonly applied in the presurgical evaluation of epilepsy, comparing a single patient against a group of healthy controls. Martin et al, 2015[5] reviewed various studies that

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-Strasse 40, 37075 Göttingen, Germany; e-mail: raviteja.kotikalapudi@gmail.com

Indicates article with supplemental on-line tables.

Indicates article with supplemental on-line photo.

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

have reported the potential use of VBM based on T1 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)[7-9] and statistical cutoffs to elucidate VBM findings.[10,11] Studies have shown that changing smoothing kernel size[9] and statistical thresholds have a direct impact on the VBM findings,[10] also affecting the specificity of the results.[10]

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 Segment). This approach enables combining different MR image contrasts. Lindig et al, 2018[12] 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 epilepsy.

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 MR imaging negative for focal epilepsy.

MATERIALS AND METHODS

Subjects

We recruited 62 healthy volunteers (N_C; 36 women, 26 men; mean age, 27.5 years) and 13 patients with MR imaging negative for cryptogenic epilepsy (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-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[12] and consists of 0.9-mm^3 isotropic 3D T1-weighted MPRAGE (= T1), 3D T2-weighted sampling perfection with application-optimized contrasts by using different flip angle evolution (SPACE sequence; Siemens) (= T2), and 3D inversion recovery prepared SPACE sequence; Siemens) (= FLAIR).

Image Processing

At first, all scans were converted from DICOM to NIfTI format using MRIConvert (https://lcni.uoregon.edu/downloads/mriconvert). The postprocessing was performed with SPM12 in a Matlab R2016a (MathWorks, Natick, Massachusetts) environment and 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^3.[13] 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]) (Good et al, 2001). 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).[14,15] 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 nonepileptogenic 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.[8] 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 1*D*) 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 ED 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 nonepileptogenic 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 1*A*, -*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 1*C*). 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 8.5%, with specificity between 9.7% and 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 are provided in Table 2.

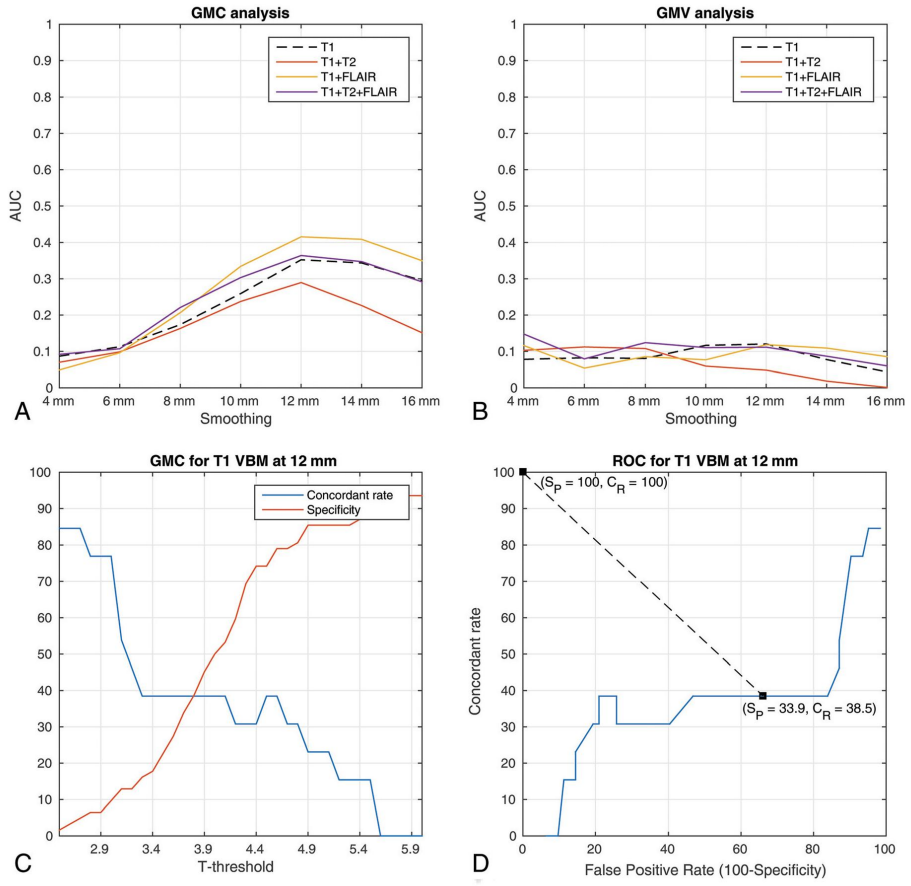


FIG 1. Diagnostic performance of different VBM models. *A* and *B*, The area under the curve for different smoothing kernels for gray matter concentration and volume. *C*, An example of an intersection plot for T1 VBM for the concordant rate and specificity against statistical cutoffs. *D*, An example 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[a]

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

[a] 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.

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 comparison 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%–16.1% of findings were marked as nonepileptogenic, while 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%.

Table 2: GMV analysis for VBM models at 12 mm and a T-threshold of 3.0[a]

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

[a] 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[a]

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	–
T1+T2	15.4	0	15.4	7.7	15.4	–
T1+FLAIR	15.4	30.8	46.2	7.7	23.1	–
T1+T2+FLAIR	15.4	15.4	30.8	7.7	23.1	–
Discordant lobe (%)						
T1	7.7	38.5	46.2	7.7	30.8	15.4
T1+T2	7.7	23.1	30.8	–	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	–	30.8	7.7

[a] 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[a]

	All lobes (%)			
Models	Nonepileptogenic	Unclear	Artifacts	Corrected Specificity (Excluding Artifacts and Nonepileptogenic)
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

[a] 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).

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[7-9,16-18] 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)[10] as well as comparisons at multiple statistical cutoffs.[8] We provide a systematic comparison of the 2 main parameters focused on patients with MR imaging negative 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.[19] In line with this study, we found a low specificity with decreased diagnostic performance (AUC, Fig 1*A* 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 1*C*). This finding is in agreement with a previous study on a lesional cohort based on *Z*-scores.[8] 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, accepting that there is 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 MR imaging negative for epilepsy, in which invasive diagnostics and invasive EEG are commonly performed and carry low-but-non-negligible risks for patients. Martin et al[10] 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 patients with MR imaging negative and MR imaging positive .[10] The only difference between the 2 analyses is the additional step of modulation, which is intended to preserve the original tissue volume.[6] 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). If this remains true for different MR imaging, pulse sequences with different tissue contrasts need 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.[10,16,20] 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[10] 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,[20] 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 , after visual analysis, the concordant rate for FLAIR and T1 VBM dropped from 28.6% and 14.3% to 14.3% and 0%, respectively.[21] 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.[10,20]

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 T1+FLAIR also showed a superior performance compared with T1.[12] 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.[21] Additionally, intensity-normalized FLAIR-VBM[11] 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.[10] 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).

Table 5: Nonvisible findings: visual analysis results for GMC[a]

	All Lobes (%)	
Models	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

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

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[10] showed that only normalized FLAIR-VBM had more concordant-than-discordant findings in patients with MR imaging negative 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.[22,23] 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 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 meaning. 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.

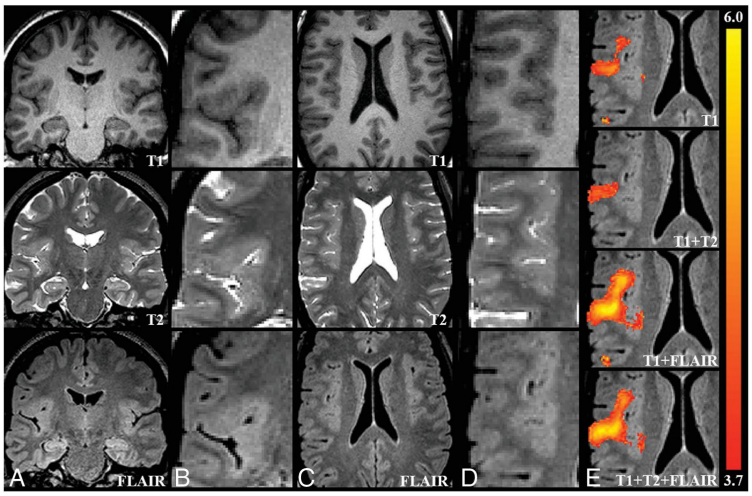


FIG 2. Sample case (patient 4). *A* and *C*, An overview in the native space T1, T2, and FLAIR images. *B* and *D*, Zoomed-in images focusing on a possible epileptogenic lesion. A subtle disruption of cortical morphology is visible in this figure, more prominent on FLAIR images. This finding is concordant 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*).

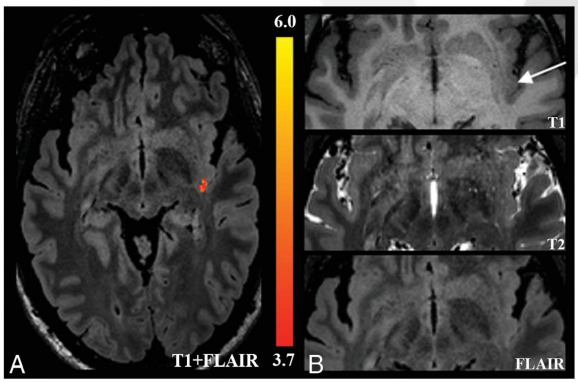


FIG 3. Sample case (patient 13). *A*, A VBM finding detected only by T1+FLAIR in the left temporoinsular region overlaid on the native space FLAIR image. *B*, Magnified images of T1, T2, and FLAIR suggesting a blurred gray-white matter junction (*arrow*). This finding is concordant with the clinical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

Diagnostic Value of Multispectral FLAIR-VBM in Focal Epilepsy

A major advantage of multispectral VBM lies in simultaneously including FLAIR along with T1. In addition, 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.[2,21] Thus, it is not surprising that FLAIR-based methods could detect those lesions better than T1 alone.[11,12] 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.[12,24,25] 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.[1] 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.

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 [formula: see text] 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

On-line Table 1: The clinical details of patients with cryptogenic epilepsy

Case No.	Age (yr)	Sex	Age of Onset (yr)	Duration of Illness (yr)	Hypothesis	PET/CT/MRI	Expert Consensus Epilepsy Surgery Case Conference	Intra-Cranial EEG	Surgery	Outcome after Surgery	Seizure-Free Duration	Histopathologic Findings	Neuropsychological Assessment	Noninvasive Video-EEG-Monitoring
1	36	F	10	25	R-F	Hypo-metabolism R-F	Concordant hypothesis, intracranial EEG indicated	R-F	Yes	Seizure-free	12 mo (Engel I)	Unremarkable CNS tissue	L/R-F	Interictal: rare sharp waves L-F/T; ictal: R-F/C
2	50	M	8	42	R-T	Unremarkable	Concordant hypothesis, intracranial EEG indicated	R-T	Yes	Not seizure-free	1 mo (Engel II)	Hippocampal sclerosis	R-F/T	Interictal: R-F and L-T; ictal: R-F/T
3	31	F	10	21	R-F/T	Unremarkable	Concordant hypothesis, intracranial EEG indicated	–	–	–	–	–	R-F/T	Interictal: spikes and focal slowing R-F/T; ictal: R-F/T
4	38	F	0	38	R-F	Unremarkable	Concordant hypothesis, intracranial EEG indicated	Multi-focal, core area R-F	–	–	–	–	R-F	Interictal: spikes R-F/C and R-F/T; ictal: R-F/C
5	55	M	38	17	L/R-T	–	Concordant hypothesis, intracranial EEG indicated	–	–	–	–	–	No localization possible	Interictal: spikes and focal slowing L/R-F/T; ictal: L/R-F/T
6	37	F	9	27	L-T/P/O	Unremarkable	Concordant hypothesis, intracranial EEG indicated	L-T/P/O	Yes	Not seizure-free	12 mo (Engel I)	Unremarkable CNS tissue	No localization possible	Interictal: spikes L-O and L-T/P/O; ictal: L-T/P/O
7	18	M	17	1	L-F	–	Concordant hypothesis due to eventual medical seizure freedom, no further diagnostics indicated	–	–	–	–	–	–	Interictal: spikes L-F/C; ictal: L-F/C
8	24	M	20	4	L-F/T	–	Concordant hypothesis due to eventual medical seizure freedom, no further diagnostics indicated	–	–	–	–	–	L-F/T	Interictal: unremarkable; ictal: L-F/T
9	33	M	15	18	L-T	–	Concordant hypothesis due to eventual medical seizure freedom, no further diagnostics indicated	–	–	–	–	–	No localization possible	Interictal: spikes and slowing L-F/T; ictal: L-F/T
10	42	F	8	34	R-F/T	–	Concordant hypothesis due to medical seizure reduction, no further diagnostics indicated	–	–	–	–		No localization possible	Interictal: unremarkable; ictal: R-F/T
11	27	F	10	17	L/R-T	Unremarkable	Concordant hypothesis, intracranial EEG indicated	L/R-T (bilateral)	–	–	–	–	L/R-F/T	Interictal: spikes and slowing L/R-F/T; ictal: L/R-F/T
12	33	M	18	15	L-T	–	Concordant hypothesis, intracranial EEG indicated, patient declined	Offered	–	–	–	–	L-F/T	Interictal: spikes L/R-F/T; ictal: L-F/T
13	43	M	14	29	L/R-T	–	Concordant hypothesis, intracranial EEG indicated, patient declined	Offered	–	–	–	–	L-F/T	Interictal: spikes R-T; ictal: Spikes L/R-T

Note:—L/R indicates left/right; F/T/P/O/C, frontal/temporal/parietal/occipital/central;

[a] Clinical details of the patients comprising patient No., age, sex, age on onset, duration of illness, PET/CT/MRI availability, expert consensus, availability of intracranial EEG data, surgery, outcome after surgery, seizure-free duration, histopathologic findings, neuropsychological assessment, and noninvasive video-EEG monitoring are shown.

On-line Table 2: AUC values for all VBM models at different smoothing levels[a]

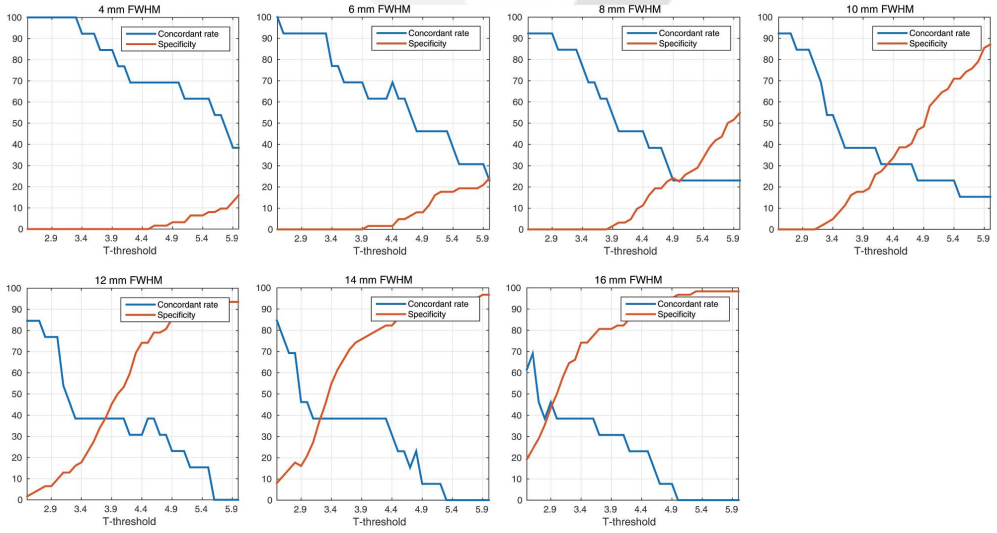
VBM	4 mm	6 mm	8 mm	10 mm	12 mm	14 mm	16 mm
GMC							
T1	0.09	0.11	0.17	0.26	0.35	0.34	0.30
T1+T2	0.07	0.10	0.16	0.24	0.29	0.23	0.15
T1+FLAIR	0.05	0.10	0.21	0.33	0.42	0.41	0.35
T1+T2+FLAIR	0.09	0.11	0.22	0.30	0.36	0.35	0.29
GMV							
T1	0.08	0.08	0.08	0.12	0.12	0.08	0.04
T1+T2	0.10	0.11	0.11	0.06	0.05	0.02	0.0
T1+FLAIR	0.12	0.05	0.09	0.8	0.12	0.11	0.09
T1+T2+FLAIR	0.15	0.08	0.12	0.11	0.11	0.09	0.06

[a] All the AUC values across statistical cutoffs (2.5–6 in a step size of 0.1) are shown for each smoothing level from 4 to 16 mm for all VBM models.

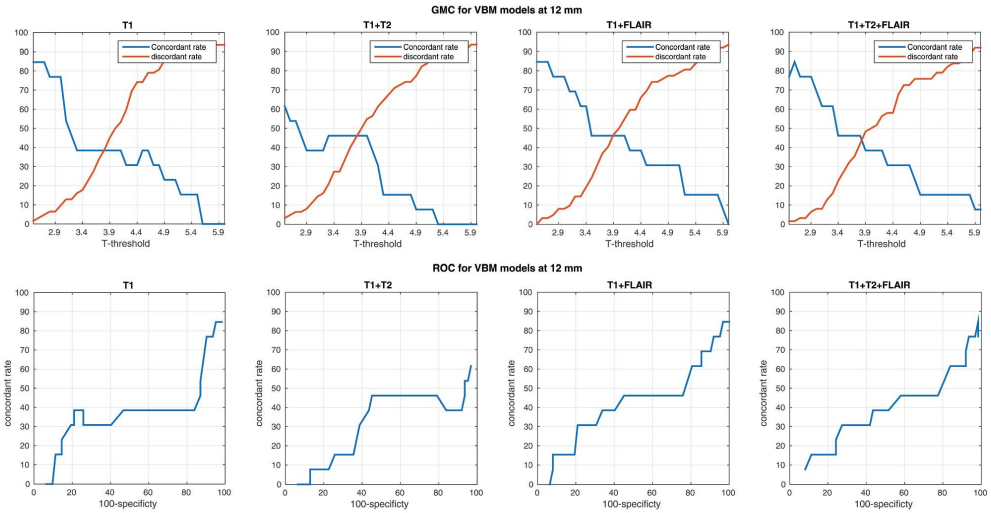
On-line Table 3: VBM findings rated as potentially epileptogenic stratified in concordant and discordant lobes for all models[a]

	Concordant Lobe				Discordant Lobe			
Case No.	T1	T1+T2	T1+FLAIR	T1+T2+FLAIR	T1	T1+T2	T1+FLAIR	T1+T2+FLAIR
3	–	–	Yes	Yes	Yes	Yes	Yes	Yes
4	Yes	Yes	Yes	Yes	Yes	–	Yes	–
5	Yes	Yes	Yes	Yes	Yes	–	–	–
6	Yes	–	Yes	Yes	–	–	–	–
7	Yes	–	Yes	–	–	–	–	–
9	–	–	–	–	–	Yes	Yes	Yes
10	–	–	–	–	Yes	–	–	–
12	–	–	–	–	Yes	Yes	Yes	Yes
13	–	–	Yes	–	Yes	Yes	Yes	Yes

[a] VBM findings in patients confirmed as potentially epileptogenic across all models for concordant and discordant lobes are presented.



ON-LINE FIG 1. Concordant rate and specificity for T1 VBM at variable statistical cutoffs and smoothing levels. Concordant rate and specificity for increasing smoothing levels from 4 to 16 mm are plotted for GMC analysis. As the smoothing level increases, the intersection point moves toward the left (lower T-threshold). FWHM indicates full width at half maximum.



ON-LINE FIG 2. Diagnostic performance for different VBM models. The plots show the concordant rate and specificity for different T-thresholds at 12-mm smoothing for all VBM models (*upper panel*). The *lower panel* shows receiver operating characteristic curves for all VBM models at 12-mm smoothing.

4.3. Study III (in preparation for submission)

Abstract

Purpose

MP2RAGE (Magnetization Prepared 2 Rapid Gradient Echo) is a new, semi-quantitative MRI sequence with good grey/white contrast and less susceptibility to image bias. Both properties could have the potential to improve lesion detection in focal epilepsy, particularly in the context of post-processing like VBM (voxel-based morphometry) in patients with a normal conventional MRI (MRI-negative). MP2RAGE is based on a modified MPRAGE sequence measured at two different inversion times (INV1 and INV2). In this study, we aim to compare the tissue segmentation quality of MP2 and its multispectral combinations; mainly MP2+INV1 and MP2+INV2 with existing T1 (MPRAGE) and T1+FLAIR approaches. Most important, we systematically evaluated the diagnostic performance at variable smoothing levels and statistical cutoffs in detecting GM structural abnormalities in previously MRI-negative focal epilepsy patients.

Patients and methods

We recruited 31 healthy controls at 21 patients with focal epilepsy (n =5 with known lesion; n = 16 with MRI-negative focal epilepsy diagnosis). 3D T1-, T2-weighted FLAIR and MP2RAGE images were acquired on Magnetom Prisma, Siemens 3T scanner, with a 64 head-channel and 1mm^3 of voxel-resolution. VBM was performed for single-channel T1, MP2 and for multispectral combinations. Effects of tissue segmentation for GM (gray matter), WM (white matter) and CSF (cerebrospinal fluid) were analyzed amongst all combinations in healthy controls. Individual patient abnormalities were analyzed across single-channel and multispectral VBM variants, at variable smoothing levels (4mm to 16mm, step size = 2) and statistical cutoffs (2.5 to 6, step-size = 0.1).

Results

We found that MP2 alone gave better delineation of deep, subcortical nuclei but was prone to misclassification of dura/vessels as grey matter, even more than with conventional T1. The addition of multispectral combinations (INV1, INV2 or FLAIR) could markedly reduce this effect. MP2+INV1 yielded generally thinner GM segmentation allowing better differentiation of white matter and neighboring gyri. For the lesion detection in MRI-positive epilepsy patients, all the VBM models showed good sensitivity between 80-100%, except for INV1+INV2 (60%). In the most challenging previously MRI-negative cases, MP2+INV1 was found as the best amongst all models with a concordant rate of 37.5%, specificity of 51.6% and concordant to discordant ratio of 0.60 at a smoothing level of 14mm and statistical cutoff of 3.3. This was generally better than T1 alone (concordant rate 18.8%, specificity 35.5%, concordant to discordant ratio 0.27) and T1+FLAIR (concordant rate 25.0%, specificity 35.5%, concordant to discordant ratio 0.57).

Conclusion

We found that VBM of MP2RAGE, especially with multispectral combinations of INV1 and INV2, have tissue segmentation advantages over conventional T1-based methods. MP2+INV1 VBM performs superior to conventional VBM variants in detecting epileptogenic abnormalities in previously MRI-negative patients, even compared to the currently best standard of T1+FLAIR. Additionally, MP2RAGE has the advantage to be available at magnetic field strengths ≥3T.

Keywords

MP2RAGE, VBM, segmentation, focal epilepsy, focal cortical dysplasia, MRI-negative, smoothing, statistical cutoff, image processing

Introduction

Focal epilepsy constitutes approximately 60% of all epilepsies (Rosenow and Lüders, 2001), and is characterized by 'focal onset', i.e., seizures originating in a local brain region (Scheffer et al., 2017). About 30% of focal epilepsy patients experience disabling seizures medically refractory/resistant to anticonvulsants (Kwan and Brodie, 2000). For these patients epilepsy surgery can be very beneficial, provided a resectable, focal area of seizure onset is identifiable (Wiebe et al., 2001). Before surgery in these patients, a presurgical evaluation is conducted, which includes brain MRI as important modality (Duncan, 1997). Cortical malformations, mainly focal cortical dysplasias (FCD), are one of the main causes associated with refractory focal epilepsy that are identifiable in MRI. Typical imaging features are blurred gray-white matter (GM-WM) junctions, cortical thinning/thickening and hypo- or hyperintense MR signal (Blumcke et al., 2011). However, a relevant proportion of lesions escape visual detection. Approximately half (30-50%) of patients undergoing surgery without MRI visible lesion eventually have cortical dysplasia / FCD upon histological investigations (Bernasconi et al., 2011; Wang et al., 2013). Potential reasons for missing an FCD are their subtlety and, at times, small size as well as variable location. Hence, such scans are mostly considered as normal MRI (MRI-negative patients). Consequently, patients have worse surgical outcomes compared to cases with visibly known/proven lesions on MRI (Tellez-Zenteno et al., 2010). In those patients, neuro-computational methods are shown to be helpful in detecting subtle lesions, previously missed upon visual analysis (Kotikalapudi et al., 2018; Martin et al., 2015; Martin et al., 2017).

Voxel-based morphometry (VBM) is a well established method for computer-aided, volumetric MRI processing (Ashburner and Friston, 2000). Usually applied on T1 images (T1 VBM), it can successfully detect local GM concentration/volume changes in patients with lesional focal epilepsy (Bonilha et al., 2006b; Colliot et al., 2006; Martin et al., 2015). VBM is based on tissue segmentation, where the brain is classified into three main tissue categories, namely: gray matter (GM), white matter (WM) and

cerebrospinal fluid (CSF) (Ashburner and Friston, 2005). MRI technical factors (e.g. image bias) directly impact on the quality of tissue segmentation and influence the VBM process (Focke et al., 2011). Improved tissue segmentation can be achieved through combining multimodal (-spectral) MRI contrasts in the segmentation process (Alfano et al., 1997; Ashburner and Friston, 1997; Ashburner and Friston, 2005; Fletcher et al., 1993; Lambert et al., 2013; Vannier et al., 1985). Recent studies confirmed superiority of multispectral over conventional T1-onls segmentation by addressing issues such as overestimation of dura and vessels as GM and also improve cortical GM segmentation (Lindig et al., 2017; Viviani et al., 2017a; Viviani et al., 2017b). With the addition of T2/T2-FLAIR weighted, VBM variants (T1+FLAIR, T1+T2) have also shown improved lesion detection over existing T1 VBM, in both lesional and non-lesional MRI focal epilepsy (Kotikalapudi et al., 2018; Lindig et al., 2017).

Particularly at higher field strengths of ≥ 3T, image bias due to static magnetic (B0) and radio-frequency field (transmission; $B1^+$ and reception; $B1^-$) inhomogeneities is problematic for segmentation algorithms (Focke et al., 2011). Marques et al., addressed this issue by developing MP2RAGE (Magnetization Prepared 2 Rapid Acquisition Gradient-Echo) sequence, which yields T1-weighted images that are, at least partially, corrected for image bias intrinsically (Marques et al., 2010). This is achieved by combining two images acquired at different inversion times (TI1; heavily T1-weighted or INV1, TI2; heavily proton density-weighted or INV2), interleaved in one sequence to calculate the MP2RAGE. The resultant images show enhanced contrast-to-noise ratio; (especially GM-WM contrast), independent of T_2^*, proton density, $B1^-$ and reduced $B1^+$ inhomogeneities. Newer studies have shown promise of MP2RAGE in improving visualization of lesions (Beck et al., 2018; Pittau et al., 2018). Moreover, reduced intensity inhomogenities should also improve tissue segmentation which facilitates VBM analysis for lesion detection (Ashburner and Friston, 2005). However, MP2RAGE and the multispectral MP2RAGE variants have not been systematically analyzed for detecting subtle epileptogenic lesions in a VBM approach. Moreover, it was previously shown that the performance of lesion detections depend critically on the choice of smoothing and statistical thresholds or t-scores (Kotikalapudi et al., 2018; Martin et al., 2017). The ideal parameters for an MP2RAGE VBM are yet unclear.

Hence, the major aim of this study is to explore applications of MP2RAGE based VBM analyses in previously MRI negative focal epilepsy patients. To achieve this, at first we systematically compared the different MP2RAGE based multispectral combinations with existing T1 and T1+FLAIR approaches, in a healthy control cohort. Subsequently, we analyzed the diagnostic performance of single-channel T1, MP2RAGE and multispectral variants at multiple smoothing levels and statistical thresholds, in patients and controls. These results should clarify if MP2RAGE VBM is capable of detecting epileptogenic lesions in focal epilepsy patients with negative conventional MRI and how this approach would compare to conventional T1 and T1+FLAIR VBM.

Methods

Data acquisition

We recruited 31 healthy controls (N_C; 14 females, mean age=28.4) and 21 patients (N_P; 7 females, mean age=31.5) with focal epilepsy. Five of the focal epilepsy patients had a visible lesion concordant with malformation of cortical development (MCD). Sixteen patients were labelled as MRI-negative on expert neuroradiology review. All patients had undergone comprehensive presurgical diagnostics including non-invasive video-EEG telemetry, epilepsy-dedicated 3T-MRI and neuropsychological assessment.

Three patients in MCD group have undergone invasive EEG and three patients have been surgically resected with histology showing evidence for FCD Type II b in two patients. For the MRI-negative patient group, invasive EEG was performed in three patients. So far, None in this group have been surgically resected. Clinical details of patients are summarized in supplementary table 1. A clinical lobar hypothesis of epilepsy onset was derived through expert consensus in the epilepsy case conference with all available information in the presurgical epilepsy program. VBM findings analyzed in this study were not considered for forming the clinical hypothesis.

Imaging data was acquired on 3T Magnetom Prisma, Siemens MRI at the University Hospital Tübingen, Germany. The acquisition protocol consisted of 3D T1-weighted Magnetization-Prepared Rapid Gradient-Echo; MPRAGE (TE=2.98ms, TR=2300ms, TI=900, flip angle=9°), a 3D T2-weighted Sampling Perfection with Application optimized Contrasts using different flip angle Evolution-Fluid-Attenuated Inversion Recovery; SPACE-FLAIR (TE=388ms, TR=5000ms, TI=1800ms, flip angle=120°) and a 3D T1-weighted Magnetization-Prepared 2 Rapid Gradient-Echo; MP2RAGE; (INV1; TI1=700 ms, flip angle=4° and INV2; TI2=2500ms, flip angle=5°; TE=2.98ms, TR=5000ms) using a 64 channel head coil with an isotropic resolution of $1mm^3$. Subjects were considered for image post-processing steps only when T1, FLAIR and MP2RAGE images were available and did not have relevant artifacts (motion in particular) in the visual quality control. All datasets/subjects passed this condition.

Image processing

All DICOM images were converted to NIFTI file format using mriconvert (http://www.lcni.uoregon.edu/~jolinda/MRIConvert). SPM12 (www.fil.ion.ucl.ac.uk) based on MATLAB R2016a (The Math Works, Natick, MA) was used for image processing. MP2 images were reconstructed and readily available from the scanner, whish uses the following equation for reconstruction:

$$\mathrm{MP2} = \frac{\mathrm{INV1}^{*}\,\mathrm{INV2}}{|\mathrm{INV1}|^2 + |\mathrm{INV2}|^2}$$

Detailed mathematical calculations and equations for MP2 image reconstruction are explained in other studies (Marques et al., 2010).

VBM was performed for single-channel T1, MP2 and multispectral combinations, namely: T1+FLAIR, MP2+FLAIR, MP2+INV1, MP2+INV2 and INV1+INV2. For multispectral combinations, co-registration was performed using a normalized mutual cost function. For co-registration, reference images were T1 (in T1+FLAIR), MP2 (in MP2+FLAIR, MP2+INV1 and MP2+INV2), and INV1 (in INV1+INV2), whereas the latter images in the combinations served as the source images. Unified segmentation (Ashburner and Friston, 2005) was applied with default settings of bias regularization; 0.0001 and bias cutoff; FWHM 60mm. Next, gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) tissue probability maps were spatially normalized with an isotropic resolution of $1mm^3$ using DARTEL, based on the respective native space GM and WM maps and building a custom template in MNI space for each multispectral combination (Ashburner, 2007). During the normalization process, images were modulated to preserve tissue quantity (in case of group level analysis) and unmodulated images to preserve tissue concentration (in case of individual subject versus group analysis) (Good et al., 2001). As a final step, images were smoothed using Gaussian kernel sizes; 4mm to 16mm (step size of 2mm) full width at half maximum (FWHM).

Comparisons for absolute tissue volumes

Native segmented GM, WM and CSF were used to calculate the tissue volumes for healthy controls across segmentation models. For this purpose, voxel values (ranging from 0 to 1) were summed and multiple with the voxel volume ($1mm^3$) to yield a tissue class specific volume. To estimate the total intracranial volume (TIV), the absolute volumes of GM, WM and CSF were added. To assess significant differences in segmentation models, one-way repeated measures ANOVA with adjustment for multiple comparison(Bonferroni), was conducted in SPSS (IBM SPSS Statistics 22), similar to our previous study on multispectral segmentation (Lindig et al., 2017).

Voxel-based comparison of multispectral variants

Smooth normalize modulated GM, WM and CSF images of healthy controls were used to perform group comparisons across different segmentation models. T1+FLAIR segmentation was used as a reference, since this had the best overall segmentation quality in our previous work (Lindig et al., 2017). Each multispectral combination was then compared against T1+FLAIR using a paired t-test in SPM, at 4mm smoothing and a statistical threshold of $p<0.05$ FWE; (family wise error).

In the individual analysis, we compared each patient against all controls (patient comparison) and each control against the rest of the controls (after removing the control in question; control comparison) in

SPM12 using statistical cutoffs from 2.5-6 (step size of 0.1). This analysis was repeated for all smoothing levels (4mm-16mm; step size of 2).

Automated individual lobar analysis

We used the MNI structural lobar atlas provided with FSL version 5.0 (Collins et al., 1995; Mazziotta et al., 2001), for automated analysis. The atlas comprised of bilateral lobes for frontal, parietal, temporal and occipital. We extracted lobar regions-of-interest from the atlas, which were the lobes of clinical hypothesis (please refer hypothesis lobes from supplementary table 2), for each patient. These lobes were considered as 'concordant lobes', i.e. concordant to the clinical hypothesis, while non-concordant lobes (remaining lobes) were labelled 'discordant lobes' individually for each patient. Since we do not expect controls to have epileptogenic findings, all lobes (bilateral: frontal, temporal, parietal and occipital lobes) were defined as discordant lobes, in case of all controls.

Performance parameters

Concordant rate; C_R (or discordant rate; D_R) was calculated for patients as the percentage ratio of number of patients with at least 1/3rd of voxels in a VBM cluster overlapping with concordant lobe (or discordant lobe) to the total number of patients;

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

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

Specificity was defined as the percentage ratio of controls with no VBM findings to total number of controls (findings with less than 1/3rd of voxels inside the brain were considered as no finding);

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

where, N_{CON} represents number of patients with concordant findings, N_{DIS} refers to number of patients with discordant findings and N_{NF} refers to number of controls with no findings. C_R, D_R and S_P were calculated for each smoothing levels; 4mm-16mm, across statistical cutoffs; 2.5-6, for all VBM models.

Estimating smoothing and T-threshold

To estimate the ideal smoothing level, we considered the single-channel MP2 VBM findings as our point of reference. Further, for every smoothing level, receiver operating characteristic (ROC) curve plots were generated with C_R and 100-S_P (false positive rate; FPR) values across statistical cutoffs. Later, the area under the ROC curve (AUC) was calculated using a trapezoidal integration function in MATLAB. The smoothing level resulting in the highest AUC in single-channel MP2 VBM was subsequently used as the reference to determine the optimal T-threshold; a statistical cutoff value giving a balanced trade-off between C_R and S_P. For this purpose, C_R and S_P values (at the selected smoothing level) were plotted for statistical cutoffs from T>2.5-6. The cutoff where C_R-S_P was least, but still

positive, was considered as the optimal T-threshold. Incorporating the estimated smoothing and T-threshold, C_R, D_R and S_P values were reported for VBM variants. Finally, the Euclidean distance (ED) was calculated as the distance between (C_R, S_P) and C_R=100, S_P=100 i.e., (100, 100);

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

The VBM variant with minimum ED was considered as the best VBM method. To further validate the selected smoothing level and T-threshold, we additionally applied these parameters to patients with visible epileptogenic lesions (suspected MCDs; n = 5).

Visual Interpretation

Visual cross-verification of VBM findings was done by one certified neuroradiologist with 10 years of experience. For group level analysis, structural differences were visually interpreted on native space across individual subjects, to assess if the VBM differences were evident on the native tissue maps. For individual level analysis, at the estimated smoothing level and T-threshold for each subject, the VBM findings across models were combined and inverse transformed to native space using the deformation utility in SPM12. These native combined VBM findings were smoothed with 1mm Gaussian kernel and were provided to the reviewer for scoring overlayed on T1, FLAIR and MP2 images, using in-house software written in MATLAB. The reviewer was blinded towards the clinical hypothesis. The reviewer visually interpreted the patient VBM findings and gave scores from 1: visible and potentially epileptogenic, 2: non-visible but potentially epileptogenic, 3: visible and likely non-epileptogenic, 4: unclear/non-visible or 5: artifact. For controls, each finding was rated as either 1: visible and likely non-epileptogenic, 2: unclear/non-visible or 3: artifact. The non-epileptogenic label was given when finding was visible on one/more structural images but was not likely related to epilepsy (e.g. perivascular spaces or microangiopathy). Unclear labels were used whenever the findings were not clearly visible to be confirmed as either epileptogenic, non-epileptogenic or as an artifact.

Results

Group level differences across segmentation models

We found significant differences among segmentation combinations for absolute volumes of GM, WM, CSF and TIV (figure 1), that were significantly different across segmentation methods (one-way repeated measures ANOVA); GM ($p<0.05$), WM ($p < 0.05$), CSF ($p<0.05$) and TIV ($p < 0.05$) (supplementary table 2). The voxel-based group level comparisons in healthy subjects using a paired t-test ($p<0.05$ FWE) for different VBM combinations are presented in figure 2 (and compare supplementary figures 1-5). These results were also qualitatively confirmed through visual inspection of individual subject segmentations in native space (figure 3, supplementary figures 6-8). First, upon visual inspection on native images, we found that the meninges and venous sinuses (straight, transverse and sigmoid) were mostly overestimated as GM by T1, MP2 and INV1+INV2 (figure 3A). Additionally, for single-channel MP2, there was also misclassification of meninges to GM in the frontal pole, superior,

middle and inferior temporal gyrus (figure 3B). With the addition of FLAIR, misclassifying these structures to GM was clearly reduced. Addition of INV1 and INV2 to MP2 also resolves the misclassification of meninges to GM, most apparent in the temporal gyrus and the occipital lobe (figure 3A-B). However, MP2+FLAIR/INV1/INV2 combinations segmented meninges and vessels as CSF tissue class (supplementary figure 8). We also observed that the cortical ribbon was segmented consistently thinner in MP2+INV1 when compared to rest of approaches. In fact, MP2+INV1 showed a clearer GM-WM and GM-CSF separation upon qualitative inspection (figure 3B, supplementary figure 6) with GM probability≥0.75-1.0. (figure 3B). Furthermore, in comparison to T1+FLAIR, except MP2+INV1, all segmentations classified parts of thalamus and mainly posterior putamen as GM. But, upon visual inspection, none of the models completely segmented thalamus as GM. However, better segmentation results were achieved by MP2, MP+INV2 and INV1+INV2, also in comparison to T1 (figure 3C, supplementary figures 3-5). Furthermore, MP2 along with its combination variants from FLAIR/INV2 and INV1+INV2 displayed a better segmentation of posterior putamen as GM with probability ≥ 0.75-1.0 (figure 3C). In fact, a major difference between T1+FLAIR, MP2+FLAIR and MP2+INV2 was in subcortical segmentation of parts of basal ganglia (figure 2, figure 3 supplementary figure 1-5). Otherwise, these three combinations are mostly similar to each other. MP2+INV1 mostly misclassified these structures as WM. Also the rest of the models including T1+FLAIR misclassified portions of thalamus, putamen and majority of globus pallidus as WM. Finally, we found that INV1+INV2 yielded high GM probabilities (≥0.5-1) for the brain stem including portions of the mid brain and medulla in comparison to all segmentation approaches (figure 2, figure 3D, supplementary figure 6). Pons was partially segmented as GM by INV1+INV2. MP2 and its combination with INV2 showed increased GM volumes in the pons (figure 2) for group comparisons at ($p<0.05$ FWE). However, visual analysis revealed that a majority of brainstem was classified as WM by all other approaches (figure 3D).

Individual VBM based on MP2 and multispectral variants in focal epilepsy

The best performing smoothing level with single-channel MP2 VBM as the reference method was found at 14mm, (AUC = 0.24, figure 4A, 4C and supplementary table 3). At this smoothing level, MP2+INV1 had the highest AUC (0.38) amongst all models. For 14mm, we found the optimal T-threshold at 3.3 (figure 4B).

Results at smoothing: 14mm and T-threshold: 3.3

While incorporating the spatial and statistical parameters, we validated the generated smoothing and statistical cutoff using patients with known lesions. For 14mm smoothing and 3.3 T-threshold, we found a sensitivity between 60-100% across all models (table 4), where MP2+FLAIR was highest with

100%, INV1+INV2 lowest with 60%. The results of automated VBM and visual analysis in MRI-negative cohort are summarized in tables 1-3. In the automated, lobar-based VBM analysis, MP2+INV1 was the best VBM variant with ED = 66.4 and C_R = 62.5, S_P = 45.2 and C_R/D_R = 0.67 (figure 4D, table 1). When visually verifying the VBM findings, MP2+INV1 (ED = 79.0) was also found to be the best diagnostic model, with C_R = 37.5, S_P = 51.6 and C_R/D_R = 0.60 (table 1, more details in table 2). The corrected specificity after visual analysis in controls increased from 19.4-45.2 to 35.5-54.8% (table 1, more details in table 3). However, concordant rate decreased from 31.3-68.8% to 6.3.-37.5% and discordant rate decreased from 56.3-93.8% to 18.8-68.8% (table 1). The percentage of non-visible findings in controls was in the range of 45.2-64.5, and in patients was 56.3-93.8% (table 3). For single-channel T1 and T1+FLAIR VBM variants, the best VBM results were found for T1+FLAIR, with ED 83.5 (T1; ED = 91.7), C_R = 68.8% (T1; 56.3%), S_P = 22.6% (T1; 19.4%) and C_R/D_R = 0.85 (0.64) (table 1). After visual inspection, T1+FLAIR was still better than T1 with ED 99.0 (T1; ED = 103.7), C_R = 25.0% (T1; 18.8%), S_P = 35.5% (T1; 35.5%) and C_R/D_R = 0.57 (0.27) (table 1, more details in table 2 and table 3).

Discussion

In this study, we have systematically compared the MP2RAGE based segmentation combinations with existing segmentation approaches; T1 and T1+FLAIR. We found that MP2 with a combination of INV1 (MP2+INV1), showed a thinner cortical ribbon segmentation with a better separation between CSF-GM and GM-WM, in comparison to rest of the single-channel and multispectral combinations. MP2 with INV2 provides better subcortical segmentation of thalamus, basal ganglia and mostly prevents misclassification of vessels and dura as GM. Most importantly, in MRI-negative focal epilepsy cohort MP2+INV1 also yielded best results with concordant rate; 37.5%, specificity; 51.6% and concordant to discordant ratio; 0.60. In the group of visibly identifiable lesions, MP2+INV1 also had a good sensitivity of 80%.

Differences among MP2 and its multispectral combinations

We found that the absolute GM volumes were significantly higher in both single-channels; T1 (55ml) and MP2 (161ml), with respect to T1+FLAIR. One important reason is the misclassification/overestimation of dura and vessel sinuses as gray matter, by single-channel segmentation methods. The overestimation is mainly due to the close proximity and similarity in signal intensity of these structures to the cortex (Lindig et al., 2017; Viviani et al., 2017a). The addition of FLAIR did resolve this misclassification both for T1 and MP2 channel. Combinations of MP2 with INV1 or INV2 also showed a similar effect of attenuating dura/vessels from GM, thus addressing this misclassification. Second, we found that MP2RAGE variant MP2+INV1 produced significantly lower amounts of GM volumes (71.9ml) in comparison to T1+FLAIR and all the other segmentation methods (96.1-233ml). Upon visual inspection, we also observe a thinner but more accurate GM segmentation

using MP2+INV1 for the cortical ribbon, in comparison to the rest (figure 3A-B). One reason for this effect can be a clear cortical delineation exhibited by INV1 images (figure 3B), in comparison to rest of the sequences. The clearer cortical delineation can be due to the consistent dark rim separating the probable WM from the cortex in INV1 images (figure 3B). A recent study has shown similar dark rim effect via null point imaging; NPI preferably using TI of 500ms (Pitiot et al., 2007). These dark rim voxels exhibit partial volume effects, where GM and WM have similar modulus for longitudinal magnetization, but in different directions, hence creating this dark ribbon. Hence, including INV1 simultaneously with MP2 improves the segmentation of gyrus surrounding the WM. Third, in comparison to all segmentation combinations, MP2 and MP2 combinations; MP2+INV2 and MP2+FLAIR, also INV1+INV2 showed a superior subcortical segmentation of thalamus and posterior putamen. This is in line with a recent study where MP2RAGE yielded greater reproducibility of GM in thalamus and putamen, when compared with T1 (Okubo et al., 2016; Streitbürger et al., 2014), attributed to better contrast and signal to noise ratio for MP2RAGE in the subcortical structures. This is mainly responsible for significantly higher GM volumes (24.2-31.1 ml) in MP2+INV2 and MP2+FLAIR in comparison with to T1+FLAIR.

VBM in focal epilepsy

Smoothing and statistical cutoffs

We have previously shown that the selection of smoothing (12mm with T1 as reference) as well as statistical threshold (t-score = 3.7) affected the detection rates in MRI negative cohort (Kotikalapudi et al., 2018). Our results suggest that a smoothing of 14mm at a liberal T-threshold of 3.3, gave the optimal trade-off between specificity and concordant rate for the MP2 contrast. In this case, these parameters can differ for different VBM variants, but to facilitate a systematic comparison between models, MP2 was taken as a reference point. This is in line with our previous study, where increased smoothing level is compensated with decreased T-threshold (Kotikalapudi et al., 2018). The worst performance was found at 4mm smoothing, reflected by minimal AUC across all models. This is also line with a previous study on single patient comparisons, which shows that reducing kernel size to 4mm or 8mm reduces experimental design robustness and results are prone to more false positives (Kotikalapudi et al., 2018; Salmond et al., 2002). In the same study, 12mm smoothing was suggested for single patient comparisons. This is closer to our obtained smoothing of 14mm and the difference in performance from 12mm to 14mm in our study is only 0.01 in terms of AUC (supplementary table 3), for MP2. So, a smoothing of 12mm could have also been considered, but at a higher statistical cutoff of 4.2 to get comparable results (supplementary figure 9). As a step of validation for the obtained parameters, we found the expected VBM sensitivity especially for patients with visible lesions within 60-100%, which is in line with previous studies based on lesional cohorts (Lindig et al., 2017; Martin et al., 2015). Various factors such as sequences used for reference (MP2 in our study), processing steps including voxel resolution ($1mm^3$ in our study) for normalization, methods of normalization (DARTEL

in our study) and nature of disorder (focal epilepsy with subtle/occult lesions in our study) could lead to different choices of smoothing and statistical cutoffs. Nevertheless, our results could provide guidance in maximizing (concordant rate or specificity) the performance of VBM models at a range of smoothing levels and statistical cutoffs (supplementary figure 9-10).

Visual interpretation of VBM findings

It can be expected that visual interpretation improves specificity of results, by eliminating false positive findings through expert knowledge. As predicted, in our study, post the visual analysis, specificity across models was higher, but also resulting in a decreased concordant rate (please refer results, mainly table 1-3), which is in line with previous studies (Kotikalapudi et al., 2018; Martin et al., 2017; Riney et al., 2012; Wang et al., 2015). Second, the range of non-visible findings in patients was slightly higher than that of controls (table 3). To some extent, higher number of non-visible findings in patients compared to controls can also point towards a clinical relevance of these findings. Thus, further investigations in this patient cohort may be pursued. The automated VBM process also had relevant number of discordant finding; 56.3%-93.8%, which reduced to 18.8%-62.5%, after the visual analysis. The most important part of our study is addressing the concordant findings, which are supported by clinical hypothesis. However, it should be noted that clinical hypothesis was based on non-invasive data in most patients, which is limited by propagating of seizure activity (Alarcon et al., 2001; Spencer et al., 1985). Hence, the discordant findings may still hold clinical significance, though this cannot be resolved at this point due. It can also be true that some of these patients might have multi-focal lesions. Also, in line with a recent study on comparisons between T1 and T1+FLAIR VBM in non-lesional epilepsy cohort, we obtained similar results showing superiority of the T1+FLAIR over T1 VBM approach in terms of concordant rate, specificity and concordant to discordant ratios (table 1-3) (Kotikalapudi et al., 2018).

Diagnostic significance of MP2 and multispectral MP2 combinations

In a recent qualitative assessment in lesional epilepsy cases at 7T, epileptogenic characteristics (cortical thickening, cortical-subcortical atrophy and blurred GM-WM junction phenomena) were well appreciated on MP2RAGE images (6/7 cases; visual sensitivity 85.7%) (Pittau et al., 2018). This is similar to our study, where MP2 (80%) and MP2 VBM variants showed a sensitivity between 80-100% in the lesional cohort (n cases = 5). One such example case is presented with figure 5, for a patient with histopathologically proven FCD type IIb in the right frontal lobe. The patient was operated and has been seizure free for the last 2.5 years. It can be observed that all the models segment the affected area as GM (figure 5C), due to the isointensity with cortex, hence revealed as abnormal voxels deviating from the norm after a statistical analysis. However, for such lesional cases especially at ≥7T, these results may not be transferable for T1 and FLAIR VBM variants like T1, T1+FLAIR and MP2+FLAIR due to increased magnetic and radio-frequency field inhomogeneities while using T1 and FLAIR sequences.

In this case, MP2 based VBM variants hold an advantage for their applicability at higher field strengths. Another case from the MRI-negative cohort is also represented here in figure 6. This finding is supported by the expert clinical hypothesis of right temporal lobe supported by non-invasive EEG recordings and neuropsychological evaluations. Though an intracranial EEG was indicated, it has not been pursued till date. In this case, only MP2+INV1 revealed abnormal GM in the regions of right amygdala. Gray matter hypertrophy is observed in this patient in the amygdala, which is most impressive on FLAIR. However, these regions are reported as abnormal only in MP2+INV1 VBM, which can be due to a better tissue classification facilitated by the addition of INV1 to MP2. A recent study at 7T has shown that T1-weighted imaging at TI = 780ms (null point imaging; NPI) helped in improved detection of lesions lesions in multiple sclerosis patients, accompanied by a superior separation of the cortical ribbon achieved through NPI (Mougin et al., 2016). A similar effect was obtained mainly by addition T1-weighted INV1 channel to MP2, while facilitating multispectral segmentation, in our study.

One advantage of MP2+INV1/INV2 combinations is the fact that these images are acquired from the same MP2RAGE sequence. This avoids the need for co-registration and movement related displacement (Marques et al., 2010). Second, especially at higher fields of ≥3T, MP2-based combinations can decrease tissue inhomogeneities (Marques et al., 2010) in comparison with existing T1 (MPRAGE) as well as FLAIR sequences. Also to address these inhomogeneities in T1(MPRAGE), alternate VBM processing pipelines have to be used, in comparison with the standard VBM approach, which is usable for MP2RAGE (Seiger et al., 2015). Recent studies on improving contrast and signal-to-noise ratio in 3D-FLAIR at 7T and recovering regions with low signal achieve this through magnetization transfer and direct signal control (Beqiri et al., 2018; Visser et al., 2010). In such situations, a multispectral combination of MP2+FLAIR can also be useful, which has shown 100% sensitivity in lesional MRI cohort in comparison to T1+FLAIR (80%), in our study. Also, at ultra higher field strengths ≥7T, MP2+INV2 combination can be used over MP2+FLAIR and T1+FLAIR, as this combination shows similar segmentation effects like T1+FLAIR and/ MP2+FLAIR, with a sensitivity of 80% in lesional MRI cases but with a superior concordant rate, specificity and concordant to discordant ratio in comparison with both MP2+FLAIR and T1+FLAIR VBM (table 1-3).

Limitations

From the cohort of MRI-negative patients no subject has yet undergone surgery. Hence, it is not possible to histologically validate the VBM findings. However, visual inspection of all VBM findings did provide diagnostically relevant information in this most challenging epilepsy cohort. Secondly, to fully assess the diagnostic yield of newer sequences like MP2RAGE and the reproducibility of results

using computational methods, it will be important to perform larger studies in collaborations with other research centers, such as meta-analysis of epilepsy cohorts in a worldwide population (Whelan et al., 2018).

Conclusion

In this study, we have systematically compared existing single-channel and multispectral segmentation models T1 and T1+FLAIR with newer models based on MP2RAGE. Further, we have compared the performance across these models for VBM in focal epilepsy patients with a negative conventional MRI. We found that segmentation based on MP2 combinations hold different advantages for different models. A finer cortical segmentation for GM was achieved using MP2+INV1, while misclassification of meninges and vessels as gray matter (GM) was best addressed via MP2+INV2. Also, superior subcortical GM classification for thalamus and posterior putamen was achieved using MP2+INV2. For visually inspected VBM results in MRI-negative focal epilepsy, we found the best performance for MP2+INV1 in detecting GM structural abnormalities with concordant rate, specificity and concordant to discordant ratio of 37.5%, 51.6% and 0.60 respectively. These performance parameters were calculated at a smoothing level of 14mm and statistical cutoff of 3.3 (t-score). At this smoothing kernel size and statistical cutoff, a sensitivity of 60-100% was achieved across all VBM models, for known lesional MRI cases with focal epilepsy (MCD).

In conclusion we find MP2RAGE-based multispectral VBM, feasible and partially superior to the best available common VBM variant (T1+FLAIR) in these most challenging cohorts of MRI-negative focal epilepsy. Considering MP2RAGE-based VBM, it also holds an additional advantage to be applicable at ultra-high fields (≥7T and above), which shall be studied further.

References

Alarcon, G., Kissani, N., Dad, M., Elwes, R.D., Ekanayake, J., Hennessy, M.J., Koutroumanidis, M., Binnie, C.D., Polkey, C.E., 2001. Lateralizing and localizing values of ictal onset recorded on the scalp: evidence from simultaneous recordings with intracranial foramen ovale electrodes. Epilepsia 42, 1426-1437.

Alfano, B., Brunetti, A., Covelli, E.M., Quarantelli, M., Panico, M.R., Ciarmiello, A., Salvatore, M., 1997. Unsupervised, automated segmentation of the normal brain using a multispectral relaxometric magnetic resonance approach. Magnetic Resonance in Medicine 37, 84-93.

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. Neuroimage 38, 95-113.

Ashburner, J., Friston, K., 1997. Multimodal image coregistration and partitioning—a unified framework. Neuroimage 6, 209-217.

Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry--the methods. Neuroimage 11, 805-821.

Ashburner, J., Friston, K.J., 2005. Unified segmentation. Neuroimage 26, 839-851.

Beck, E., Sati, P., Sethi, V., Kober, T., Dewey, B., Bhargava, P., Nair, G., Cortese, I., Reich, D., 2018. Improved Visualization of Cortical Lesions in Multiple Sclerosis Using 7T MP2RAGE. American Journal of Neuroradiology.

Beqiri, A., Hoogduin, H., Sbrizzi, A., Hajnal, J.V., Malik, S.J., 2018. Whole-brain 3 D FLAIR at 7 T using direct signal control. Magnetic Resonance in Medicine.

Bernasconi, A., Bernasconi, N., Bernhardt, B.C., Schrader, D., 2011. Advances in MRI for'cryptogenic'epilepsies. Nature reviews neurology 7, 99.

Blumcke, I., Thom, M., Aronica, E., Armstrong, D.D., Vinters, H.V., Palmini, A., Jacques, T.S., Avanzini, G., Barkovich, A.J., Battaglia, G., Becker, A., Cepeda, C., Cendes, F., Colombo, N., Crino, P., Cross, J.H., Delalande, O., Dubeau, F., Duncan, J., Guerrini, R., Kahane, P., Mathern, G., Najm, I., Ozkara, C., Raybaud, C., Represa, A., Roper, S.N., Salamon, N., Schulze-Bonhage, A., Tassi, L., Vezzani, A., Spreafico, R., 2011. The clinicopathologic spectrum of focal cortical dysplasias: a consensus classification proposed by an ad hoc Task Force of the ILAE Diagnostic Methods Commission. Epilepsia 52, 158-174.

Bonilha, L., Montenegro, M.A., Rorden, C., Castellano, G., Guerreiro, M.M., Cendes, F., Li, L.M., 2006. Voxel-based morphometry reveals excess gray matter concentration in patients with focal cortical dysplasia. Epilepsia 47, 908-915.

Collins, D.L., Holmes, C.J., Peters, T.M., Evans, A.C., 1995. Automatic 3-D model-based neuroanatomical segmentation. Human brain mapping 3, 190-208.

Colliot, O., Bernasconi, N., Khalili, N., Antel, S.B., Naessens, V., Bernasconi, A., 2006. Individual voxel-based analysis of gray matter in focal cortical dysplasia. Neuroimage 29, 162-171.

Fletcher, L.M., Barsotti, J.B., Hornak, J.P., 1993. A multispectral analysis of brain tissues. Magnetic Resonance in Medicine 29, 623-630.

Focke, N.K., Helms, G., Kaspar, S., Diederich, C., Tóth, V., Dechent, P., Mohr, A., Paulus, W., 2011. Multi-site voxel-based morphometry—not quite there yet. Neuroimage 56, 1164-1170.

Good, C.D., Johnsrude, I.S., Ashburner, J., Henson, R.N., Friston, K.J., Frackowiak, R.S., 2001. A voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage 14, 21-36.

Kotikalapudi, R., Martin, P., Marquetand, J., Lindig, T., Bender, B., Focke, N.K., 2018. (in press) Systematic assessment of multispectral voxel-based morphomtery in previously MRI-negative focal epilepsy. American Journal of Neuroradiology.

Kwan, P., Brodie, M.J., 2000. Early identification of refractory epilepsy. New England Journal of Medicine 342, 314-319.

Lambert, C., Lutti, A., Helms, G., Frackowiak, R., Ashburner, J., 2013. Multiparametric brainstem segmentation using a modified multivariate mixture of Gaussians. NeuroImage: clinical 2, 684-694.

Lindig, T., Kotikalapudi, R., Schweikardt, D., Martin, P., Bender, F., Klose, U., Ernemann, U., Focke, N.K., Bender, B., 2017. Evaluation of multimodal segmentation based on 3D T1-, T2- and FLAIR-weighted images - the difficulty of choosing. Neuroimage.

Marques, J.P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P.F., Gruetter, R., 2010. MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. Neuroimage 49, 1271-1281.

Martin, P., Bender, B., Focke, N.K., 2015. Post-processing of structural MRI for individualized diagnostics. Quant Imaging Med Surg 5, 188-203.

Martin, P., Winston, G.P., Bartlett, P., de Tisi, J., Duncan, J.S., Focke, N.K., 2017. Voxel-based magnetic resonance image postprocessing in epilepsy. Epilepsia 58, 1653-1664.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Parsons, L., Narr, K., Kabani, N., Le Goualher, G., Boomsma, D., Cannon, T., Kawashima, R., Mazoyer, B., 2001. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philos Trans R Soc Lond B Biol Sci 356, 1293-1322.

Mougin, O., Abdel-Fahim, R., Dineen, R., Pitiot, A., Evangelou, N., Gowland, P., 2016. Imaging gray matter with concomitant null point imaging from the phase sensitive inversion recovery sequence. Magnetic Resonance in Medicine 76, 1512-1516.

Okubo, G., Okada, T., Yamamoto, A., Kanagaki, M., Fushimi, Y., Okada, T., Murata, K., Togashi, K., 2016. MP2RAGE for deep gray matter measurement of the brain: A comparative study with MPRAGE. J Magn Reson Imaging 43, 55-62.

Pitiot, A., Totman, J., Gowland, P., 2007. Null point imaging: a joint acquisition/analysis paradigm for MR classification. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 759-766.

Pittau, F., Baud, M.O., Jorge, J., Xin, L., Grouiller, F., Iannotti, G.R., Seeck, M., Lazeyras, F., Vulliémoz, S., Vargas, M.I., 2018. MP2RAGE and Susceptibility-Weighted Imaging in Lesional Epilepsy at 7T. Journal of Neuroimaging.

Riney, C.J., Chong, W.K., Clark, C.A., Cross, J.H., 2012. Voxel based morphometry of FLAIR MRI in children with intractable focal epilepsy: implications for surgical intervention. Eur J Radiol 81, 1299-1305.

Rosenow, F., Lüders, H., 2001. Presurgical evaluation of epilepsy. Brain 124, 1683-1700.

Salmond, C., Ashburner, J., Vargha-Khadem, F., Connelly, A., Gadian, D., Friston, K., 2002. Distributional assumptions in voxel-based morphometry. Neuroimage 17, 1027-1030.

Scheffer, I.E., Berkovic, S., Capovilla, G., Connolly, M.B., French, J., Guilhoto, L., Hirsch, E., Jain, S., Mathern, G.W., Moshe, S.L., Nordli, D.R., Perucca, E., Tomson, T., Wiebe, S., Zhang, Y.H., Zuberi, S.M., 2017. ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. Epilepsia 58, 512-521.

Seiger, R., Hahn, A., Hummer, A., Kranz, G.S., Ganger, S., Küblböck, M., Kraus, C., Sladky, R., Kasper, S., Windischberger, C., 2015. Voxel-based morphometry at ultra-high fields. A comparison of 7 T and 3 T MRI data. Neuroimage 113, 207-216.

Spencer, S.S., Williamson, P.D., Bridgers, S.L., Mattson, R.H., Cicchetti, D.V., Spencer, D.D., 1985. Reliability and accuracy of localization by scalp ictal EEG. Neurology 35, 1567-1575.

Streitbürger, D.-P., Pampel, A., Krueger, G., Lepsien, J., Schroeter, M.L., Mueller, K., Möller, H.E., 2014. Impact of image acquisition on voxel-based-morphometry investigations of age-related structural brain changes. Neuroimage 87, 170-182.

Tellez-Zenteno, J.F., Hernandez Ronquillo, L., Moien-Afshari, F., Wiebe, S., 2010. Surgical outcomes in lesional and non-lesional epilepsy: a systematic review and meta-analysis. Epilepsy Res 89, 310-318.

Vannier, M.W., Butterfield, R.L., Jordan, D., Murphy, W.A., Levitt, R.G., Gado, M., 1985. Multispectral analysis of magnetic resonance images. Radiology 154, 221-224.

Visser, F., Zwanenburg, J.J., Hoogduin, J.M., Luijten, P.R., 2010. High-resolution magnetization-prepared 3D-FLAIR imaging at 7.0 Tesla. Magnetic Resonance in Medicine 64, 194-202.

Viviani, R., Pracht, E.D., Brenner, D., Beschoner, P., Stingl, J.C., Stocker, T., 2017a. Multimodal MEMPRAGE, FLAIR, and R2* Segmentation to Resolve Dura and Vessels from Cortical Gray Matter. Front Neurosci 11, 258.

Viviani, R., Stocker, T., Stingl, J.C., 2017b. Multimodal FLAIR/MPRAGE segmentation of cerebral cortex and cortical myelin. Neuroimage 152, 130-141.

Wang, Z.I., Alexopoulos, A.V., Jones, S.E., Jaisani, Z., Najm, I.M., Prayson, R.A., 2013. The pathology of magnetic-resonance-imaging-negative epilepsy. Mod Pathol 26, 1051-1058.

Wang, Z.I., Jones, S.E., Jaisani, Z., Najm, I.M., Prayson, R.A., Burgess, R.C., Krishnan, B., Ristic, A., Wong, C.H., Bingaman, W., Gonzalez-Martinez, J.A., Alexopoulos, A.V., 2015. Voxel-based morphometric magnetic resonance imaging (MRI) postprocessing in MRI-negative epilepsies. Ann Neurol 77, 1060-1075.

Whelan, C.D., Altmann, A., Botía, J.A., Jahanshad, N., Hibar, D.P., Absil, J., Alhusaini, S., Alvim, M.K., Auvinen, P., Bartolini, E., 2018. Structural brain abnormalities in the common epilepsies assessed in a worldwide ENIGMA study. Brain 141, 391-408.

Wiebe, S., Blume, W.T., Girvin, J.P., Eliasziw, M., 2001. A randomized, controlled trial of surgery for temporal-lobe epilepsy. New England Journal of Medicine 345, 311-318.

Tables and figures

Table 1: Automated and visual interpretation of VBM results for MRI-negative patients and controls at T-threshold = 3.3 and smoothing = 14mm. Automated and visually interpreted VBM results for GMC analysis are present in this table at a smoothing of 14mm and T-threshold of 3.3. Shown

here are the percentage of concordant rates (C_R), discordant rates (D_R) and specificity (S_P) for each model. Euclidean distance (ED) and concordant to discordant ratios (C_R/D_R) are also present. Lowest ED refers to a model with balanced trade-off between S_P and C_R, across all models.

Table 2: Patients: Visual analysis results for GMC at T-threshold = 3.3 and smoothing = 14mm. Visual analysis results for different VBM models are present in this table for both concordant and discordant lobes. The results comprise of percentage number of potentially epileptogenic (and visible), potentially epileptogenic (and not visible), potentially epileptogenic (combined), non-epileptogenic, unclear and artifacts. Potentially epileptogenic (combined) refers to percentage of patients who had either potentially epileptogenic lesion (visible/non visible or both) overlapping with the lobe of clinical hypothesis.

Table 3: Controls: Visual analysis results for GMC at T-threshold = 3.3 and smoothing = 14mm. Shown here are VBM results for controls with percentage of non-epileptogenic, unclear, artifacts and corrected specificity in controls, for different VBM models. Also shown are percentage of controls and patients with non-visible findings. Corrected specificity refers to percentage of controls without non-epileptogenic/artifact findings. Corrected specificity can be derived as 100-unclear findings (%). Non-visible findings are across all the lobes in controls and patients. In patients, non-visible findings are percentage of epileptogenic (and non-visible) and/or unclear findings.

Table 4: Automated VBM analysis results for MCD patients. Automated VBM findings of the MRI-positive (MCD) are shown, represented by unthresholded t-scores (cluster maximum) at a smoothing of 14mm, for all VBM models. At T-threshold = 3.3, the VBM variants had a sensitivity between 60-100%. * indicates findings below T-threshold of 3.3.

Figure 1: Mean of absolute volumes of gray matter (GM), white matter (WM), cerebrospinal fluid (CSF) and total intracranial volume (TIV) for controls across all models. A box plot representation of mean absolute volumes of GM (A), WM (B), CSF (C) and TIV (D) are present in this figure, across all VBM models.

Figure 2: Group level differences for T1+FLAIR with rest of the models of GM increases and decreases for gray matter volume (modulated images) analysis. Group level comparison for T1+FLAIR with rest of the models is shown in this figure. The look-up table with Red-Yellow represents increased GM volumes in labelled models against T1+FLAIR (models > T1+FLAIR), while Blue-light blue represents decreased GM volumes (models > T1+FLAIR).

Figure 3: Segmentation results in the individual native space, with up-sample from 1mm^3 to 0.5mm^3 for appreciating segmentation differences across all combinations. A) shows native T1, MP2, INV1, INV2 and FLAIR images in sagittal view, with GM segmentation probabilities (0-1) for all combinations, i.e., T1, MP2, MP2+INV1, MP2+INV2, MP2+FLAIR, T1+FLAIR and INV1+INV2 for visual review of dura and vessels. B) shows the native images and segmentation results for the visual analysis of cortical segmentation, in the coronal plane. C) shows subcortical segmentation of thalamus and putamen (mainly posterior putamen) for all segmentation combinations in the axial view and D) shows the segmentation of brain stem.

Figure 4: Assessment of smoothing and statistical cutoffs with MP2RAGE (MP2) as reference. Performance of different smoothing levels (A), the plot for concordant rates and specificity versus increasing statistical cutoffs (B), receiver operating characteristic (ROC) curve at 14mm smoothing for MP2 (C), and automated VBM findings for all models at 14mm smoothing and 3.3 T-threshold (D) are present with this figure.

Figure 5: VBM results for all segmentation approaches; single-channel and multispectral segmentations in histopathologically proven FCD type IIb. A) native T1, MP2, INV1, INV2 and FLAIR images in axial view, B) zoom-in version of the native images at the region of interest, C) segmentation approaches; T1, MP2, MP2+INV1, MP2+INV2, MP2+FLAIR, T1+FLAIR and INV1+INV2 for the lesion, D) VBM findings after SPM GLM models for two-sample t-test for all the VBM variants (single-channel and multispectral), at smoothing of 14mm and statistical cutoff of 3.3. It can be observed that MP2 and MP2 based VBM variants also segment the affected area in the lesion as GM, with lesion appearing as hypo intense on all T1-weighted images (T1, MP2, INV1 and INV2) and hyper intense on FLAIR image.

Figure 6: VBM results for an MRI-negative patient, detected only by MP2 VBM variant; MP2+INV1, presented in the coronal view. Hypertrophy in the right amygdala region is most visible on FLAIR image. MP2+INV1 VBM detects this abnormality after individual patient versus control group comparison, at smoothing of 14mm and statistical threshold of 3.3. This finding is supported by clinical hypothesis of right temporal lobe, also supported by non-invasive EEG and neuropsychological evaluations, suspecting abnormality in the right temporal lobe.

Table 1

Automated VBM results					
Model	**Concordant rate (C_R, %)**	**Discordant rate (D_R, %)**	**Specificity (S_P, %)**	**Euclidean distance (ED)**	**C_R/D_R**
T1	56.3	87.5	19.4	91.7	0.64
MP2	37.5	68.8	32.3	92.1	0.55

T1+FLAIR	68.8	81.3	22.6	83.5	0.85
MP2+INV1	62.5	93.8	45.2	66.4	0.67
MP2+INV2	43.8	62.5	32.3	88.0	0.70
MP2+FLAIR	50.0	68.8	32.3	84.2	0.73
INV1+INV2	31.3	56.3	41.9	90.0	0.57
Visual interpretation of VBM results					
T1	18.8	68.8	35.5	103.7	0.27
MP2	18.8	43.8	45.2	98.0	0.43
T1+FLAIR	25.0	43.8	35.5	99.0	0.57
MP2+INV1	37.5	62.5	51.6	79.0	0.60
MP2+INV2	31.3	37.5	45.2	87.9	0.83
MP2+FLAIR	18.8	31.3	35.5	103.7	0.60
INV1+INV2	6.3	18.8	54.8	104.0	0.34

Table 2

			Concordant lobe			
Model	**Potentially epileptogenic (and visible, %)**	**Potentially epileptogenic (and not visible, %)**	**Potentially epileptogenic (combined, %)**	**Non-epileptogenic (%)**	**Unclear (%)**	**Artifact (%)**
T1	6.3	12.5	18.8	0	43.8	25.0
MP2	6.3	12.5	18.8	0	25.0	0
T1+FLAIR	6.3	18.8	25.0	6.3	43.8	6.3
MP2+INV1	6.3	31.3	37.5	12.5	43.8	12.5
MP2+INV2	6.3	25.0	31.3	6.3	25.0	6.3
MP2+FLAIR	6.3	12.5	18.8	6.3	31.3	6.3
INV1+INV2	0	6.3	6.3	0	25.0	6.3
			Discordant lobe			
T1	6.3	68.8	68.8	37.5	68.8	25.0
MP2	6.3	43.8	43.8	25.0	37.5	12.5
T1+FLAIR	6.3	43.8	43.8	25.0	62.5	43.8
MP2+INV1	6.3	62.5	62.5	50.0	81.3	43.8
MP2+INV2	6.3	31.3	37.5	31.3	43.8	18.8
MP2+FLAIR	6.3	31.3	31.3	25.0	50.0	31.3
INV1+INV2	0	18.8	18.8	12.5	43.8	0

Table 3

Model	**Non-epileptogenic (%)**	**Unclear (%)**	**Artifact (%)**	**Corrected specificity (%)**	**Non-visible findings (controls, %)**	**Non-visible findings (patients, %)**
T1	29.0	64.5	25.8	35.5	64.5	87.5
MP2	19.4	54.8	25.8	45.2	54.8	68.8

T1+FLAIR	29.0	64.5	29.0	35.5	64.5	87.5
MP2+INV1	6.5	48.4	9.7	51.6	48.4	93.8
MP2+INV2	25.8	54.8	22.6	45.2	54.8	68.8
MP2+FLAIR	16.1	64.5	19.4	35.5	64.5	56.3
INV1+INV2	25.8	45.2	29.0	54.8	45.2	56.3

Table 4

Model	T1	MP2	T1+FLAIR	MP2+INV1	MP2+INV2	MP2+FLAIR	INV1+INV2
Smoothing = 14mm, T-threshold = 3.3							
Case 1	7.9	10.1	11.2	9.4	9.1	11.7	8.9
Case 4	2.5*	2.5*	5.3	2.3*	4.2	4.7	3.1*
Case 16	4.2	3.8	3.9	3.9	3.9	3.9	5.3
Case 17	3.6	3.3	3.2*	3.8	3.1*	3.5	2.4*
Case 21	4.6	4.8	6.8	6.9	5.7	7.5	5.9
Sensitivity (%)	80	80	80	80	80	100	60

* indicates findings below T-threshold of 3.3.

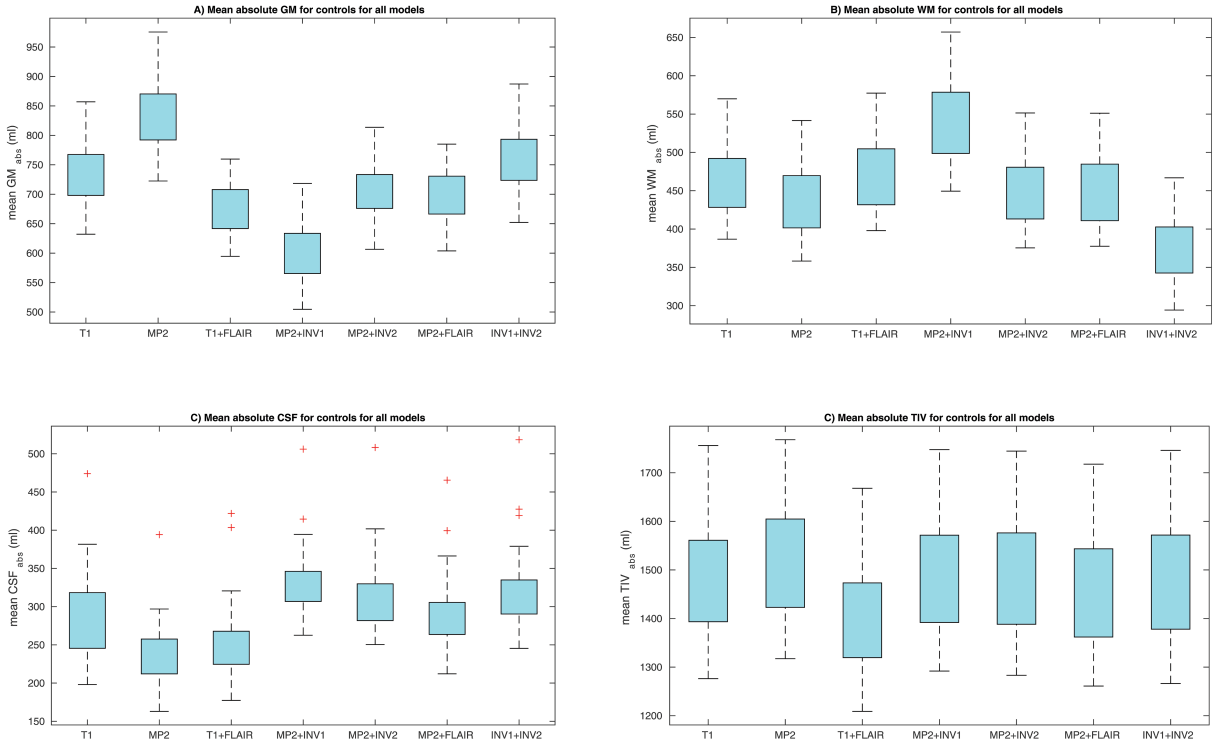


Figure 1

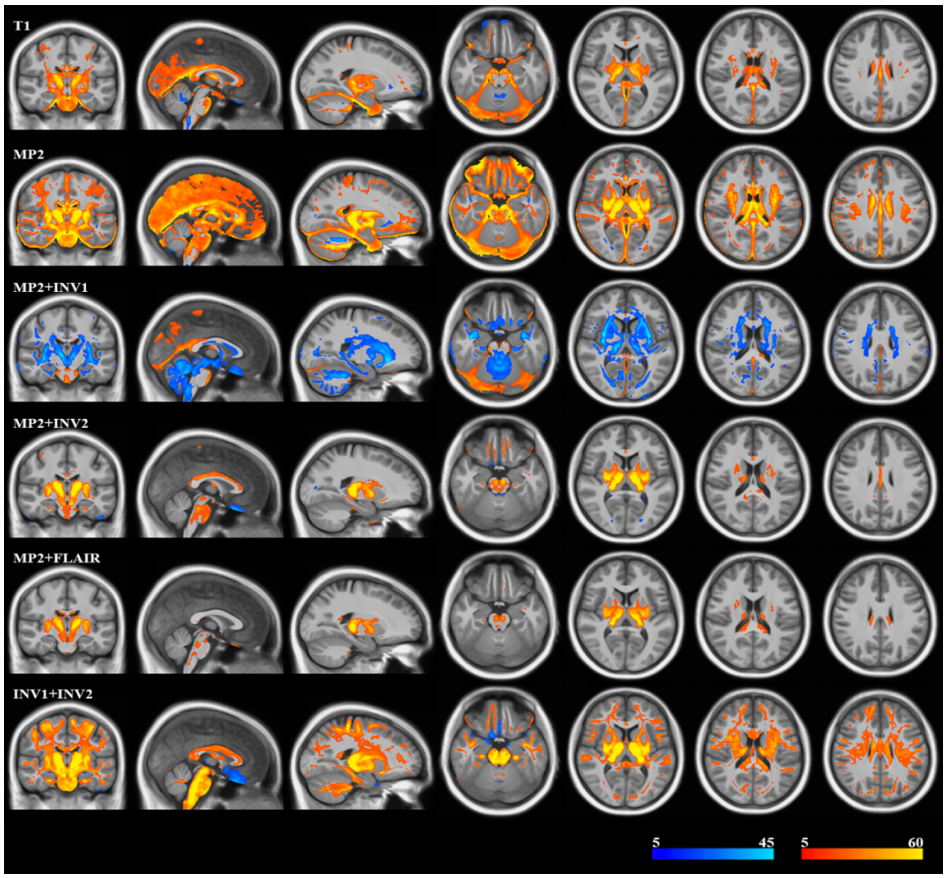


Figure 2

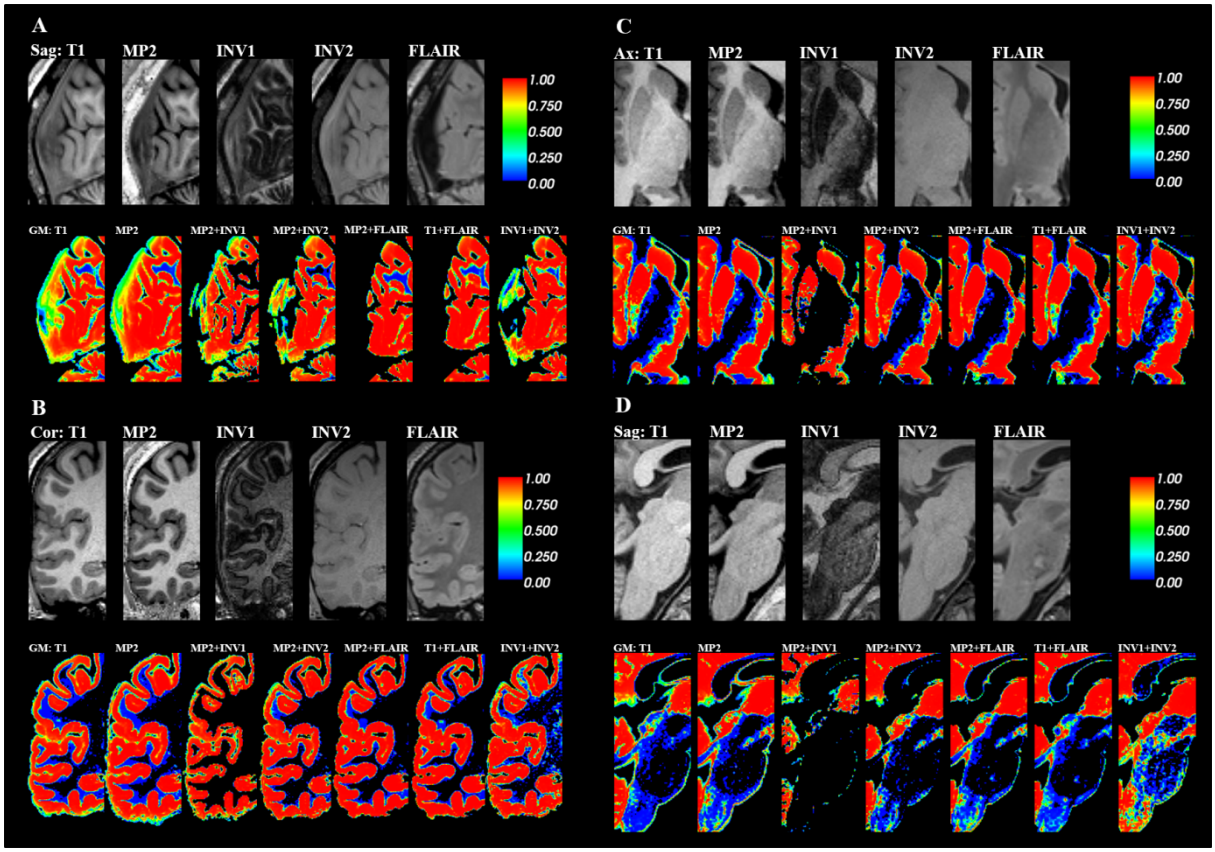


Figure 3

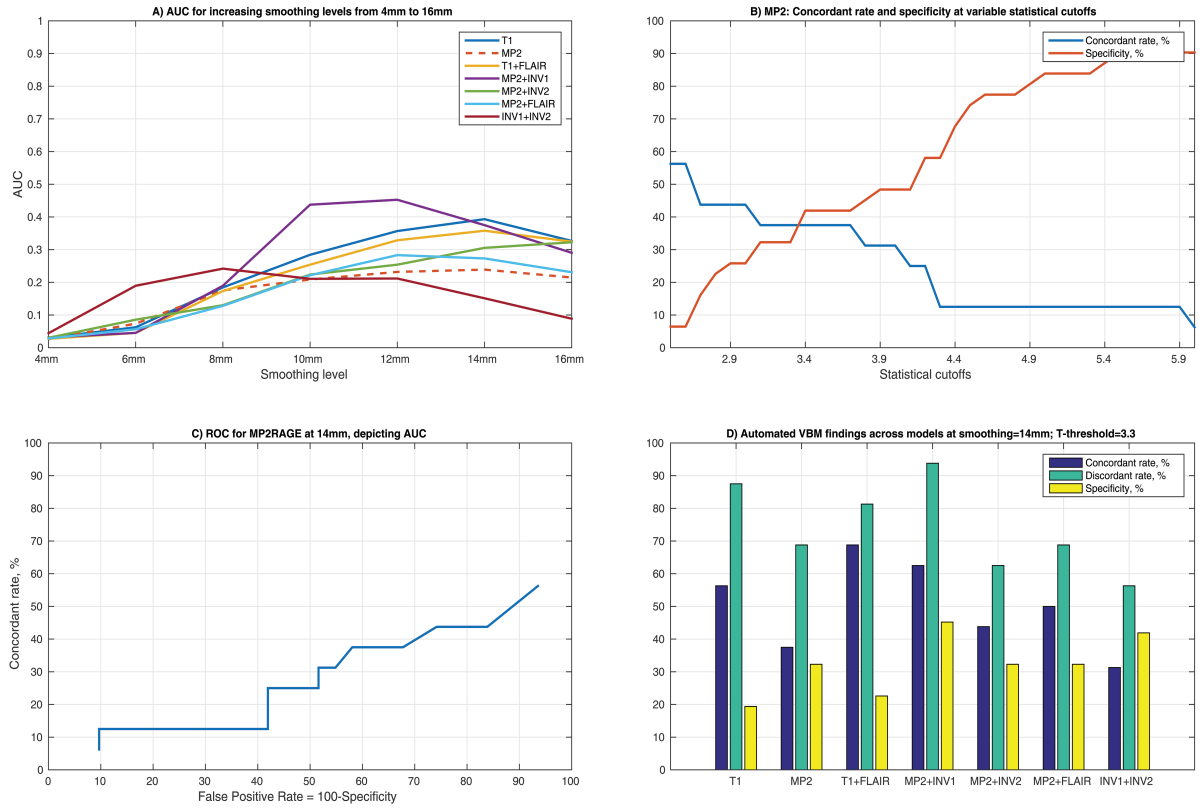

A) AUC for increasing smoothing levels from 4mm to 16mm
AUC
Smoothing level
4mm
6mm
8mm
10mm
12mm
14mm
16mm
T1
MP2
T1+FLAIR
MP2+INV1
MP2+INV2
MP2+FLAIR
INV1+INV2
B) MP2: Concordant rate and specificity at variable statistical cutoffs
Concordant rate, %
Specificity, %
Statistical cutoffs
2.9
3.4
3.9
4.4
4.9
5.4
5.9
C) ROC for MP2RAGE at 14mm, depicting AUC
Concordant rate, %
False Positive Rate = 100-Specificity
D) Automated VBM findings across models at smoothing=14mm; T-threshold=3.3
Concordant rate, %
Discordant rate, %
Specificity, %
T1
MP2
T1+FLAIR
MP2+INV1
MP2+INV2
MP2+FLAIR
INV1+INV2

Figure 4

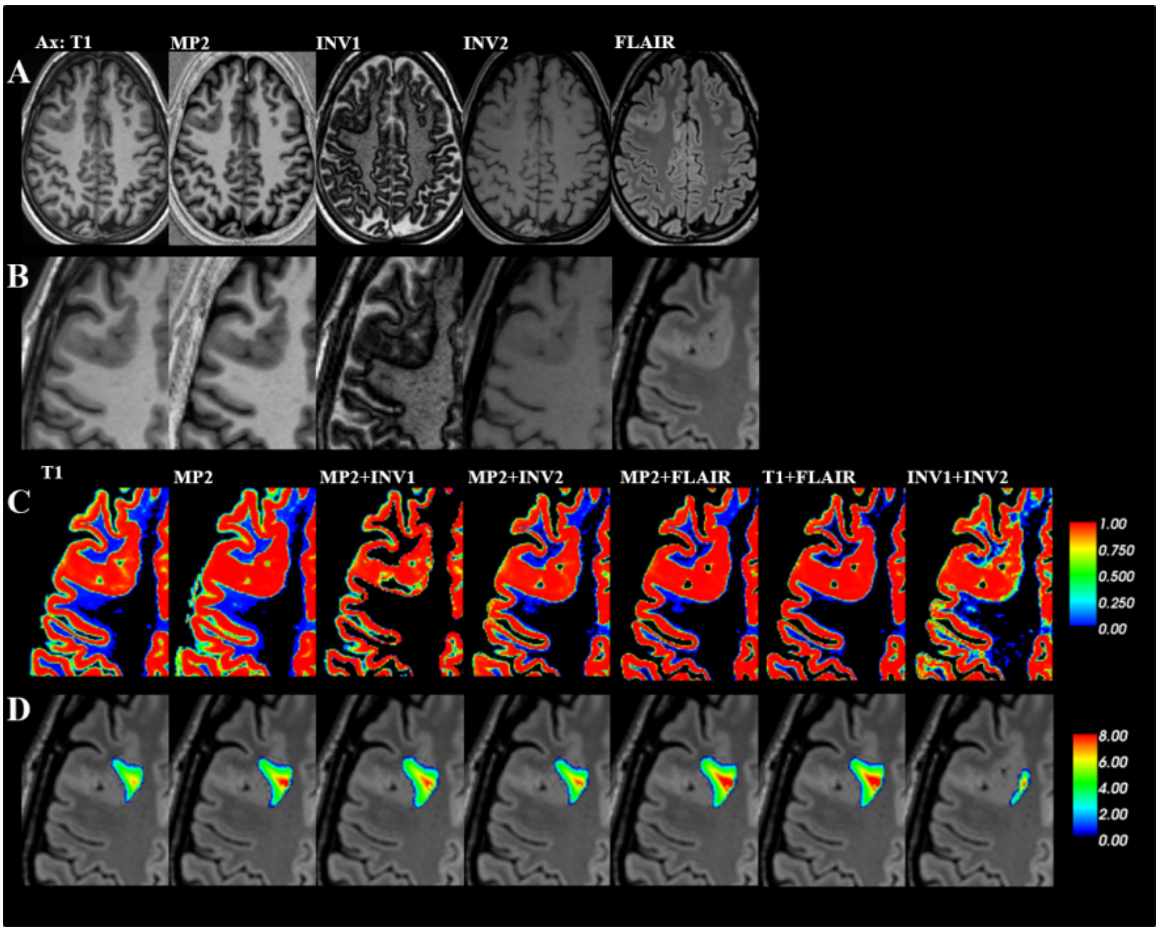

Ax: T1
MP2
INV1
INV2
FLAIR
A
B
T1
MP2
MP2+INV1
MP2+INV2
MP2+FLAIR
T1+FLAIR
INV1+INV2
C
1.00
0.750
0.500
0.250
0.00
D
8.00
6.00
4.00
2.00
0.00

Figure 5

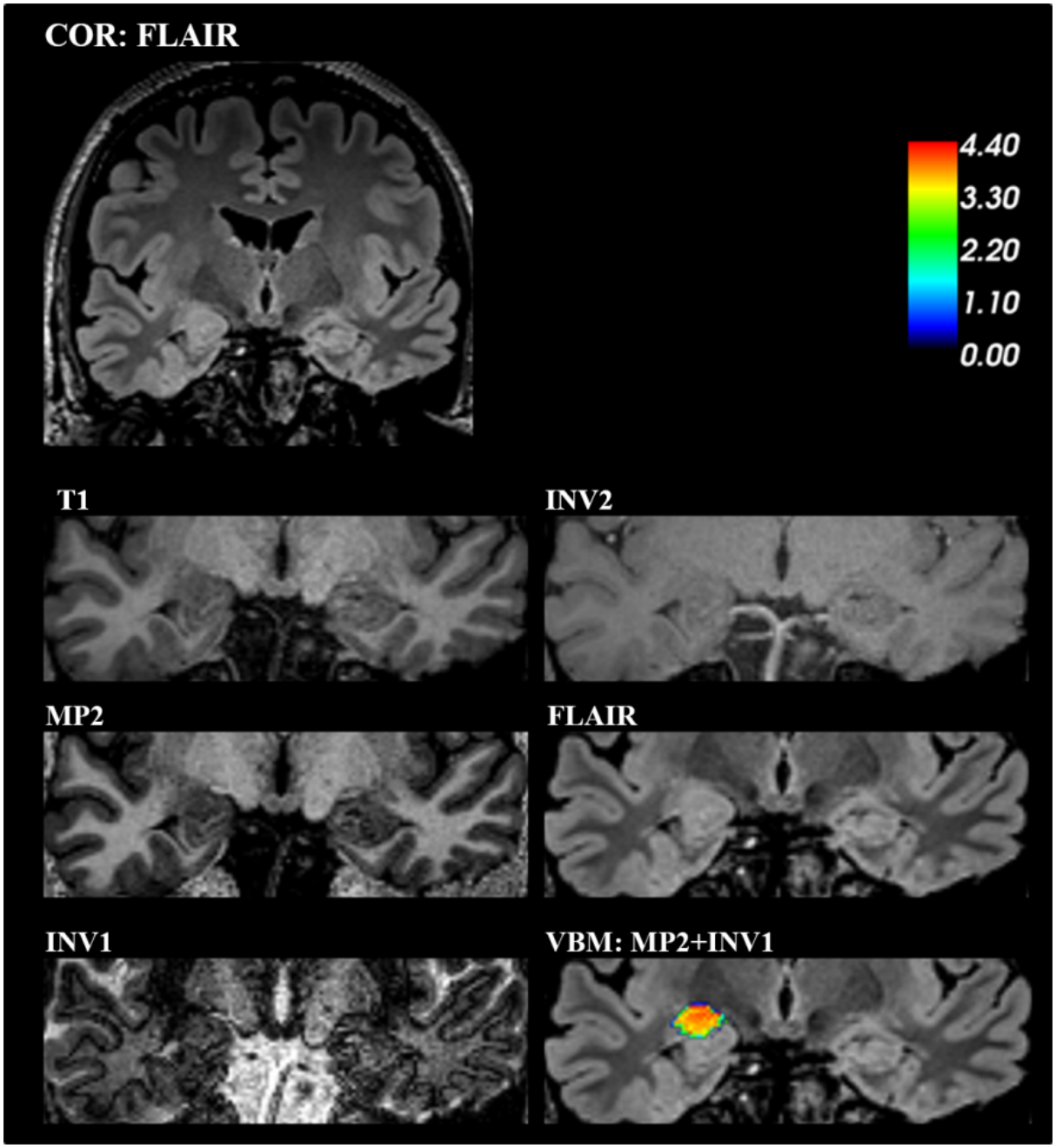


Figure 6

Supplementary table 1: Clinical details of patients with focal epilepsy.

Case ID	Age	Sex	Age of onset	Hypothesis	PET – CT/MRI when performed	Expert consensus Epilepsy surgery case conference	Intracranial EEG	Surgery	Outcome after surgery	Seizure free Duration	Histopathological findings	Neuropsychological assessment	Non-invasive Video-EEG-Monitoring
1	22	f	5	R F	not performed	Concordant hypothesis, Surgery indicated	no	Yes	Engel Class I	2.5 years	FCD IIb	F	R F
2	28	f	18	L/R T	L T	Concordant hypothesis, intracranial EEG indicated	yes	No	-	Not seizure free	-	L F-T	L/R T
3	19	m	10	R T-P-O	Unremarkable	Concordant hypothesis, intracranial EEG indicated	yes	No	-	Not seizure free	-	R T-P	R T-P-O
4	30	f	9	R I-O	not performed	Not performed	-	-	-	4 years	-	F	Unremarkable
5	19	m	4	R F-T	Unremarkable	Not performed	-	-	-	Not seizure free	-	R F-T	R F-T
6	21	m	14	L T-P	Unremarkable	Concordant hypothesis, intracranial EEG indicated	yes	no	-	Not seizure free	-	L F	L F-T-P
7	60	m	20	L/R T	not performed	Not performed	no	no	-	3 years	-	R/L F-T	L T
8	22	f	14	L F	not performed	Concordant hypothesis, intracranial EEG indicated	no	no	-	2.5 years	-	R/L F	L F
9	21	f	12	R F	not performed	Concordant hypothesis, intracranial EEG indicated	no	no	-	2 years	-	R F-T	R F
10	18	f	13	L/R F	Unremarkable	Concordant hypothesis, intracranial EEG indicated	no	no	-	1.5 years	-	R/L F	R/L F
11	28	m	26	R F-T	Unremarkable	Concordant hypothesis, intracranial EEG indicated	no	no	-	Not seizure free	-	R F-T	R F-T
12	27	m	14	L/R T	Unremarkable	Concordant hypothesis, intracranial EEG indicated	no	no	-	Not seizure free	-	R F-T	R F-T
13	22	m	16	L/R F	Unremarkable	Concordant hypothesis, intracranial EEG indicated	no	no	-	Seizure free	-	R/L F	R/L F
14	33	m	16	R T	Unremarkable	Concordant hypothesis, intracranial EEG indicated	no	no	-	Not seizure free	-	R F-T	R F-T
15	26	m	13	L F	Unremarkable	Concordant hypothesis, intracranial EEG indicated	no	no	-	3	-	R/L F	L F

16	47	m	32	L F	not performed	Concordant hypothesis, intracranial EEG indicated	yes	yes	Engel Class I	12 years	FCD IIb	L F-T	L F
17	31	f	3	L F	L F	Concordant hypothesis, intracranial EEG indicated	yes	yes	currently not classifiable	3 months	-	L F-T	L F
18	32	f	18	L T	not performed	Concordant hypothesis, intracranial EEG indicated	no	no	-	Not seizure free	-	L F-T	L T
19	51	f	30	R T	not performed	Discordant hypothesis, intracranial EEG indicated	no	no	-	Not Seizure free	-	R/L F-T	R/L F-T
20	44	m	14	L T	not performed	Concordant hypothesis, intracranial EEG indicated	no	no	-	Not seizure free	-	L F-T	R/L F-T
21	40	f	1	R F	Unremarkable	Concordant hypothesis, intracranial EEG indicated	yes	no	-	Not seizure free	-	R F-T	R F

Clinical details comprising of subject ID, age, gender, age of onset, PET-CT/MRI, clinical hypothesis, intracranial EEG, outcome of surgery, seizure freedom, histopathological findings, neuropsychological assessment and non-invasive scalp EEG information are presented with this table. This information contains both MRI-negative and MRI-visible lesional (suspected MCD) cases. Note: '-' refers to information not available/applicable.

Supplementary table 2: Results for repeated measures one-way ANOVA for in-between group differences in mean absolute volumes of gray matter (GM), white matter (WM), cerebrospinal fluid (CSF) and total intracranial volume (TIV) across different segmentation models.

GM	Mean(ml)± SE	T1	MP2	T1+FLAIR	MP2+INV1	MP2+INV2	MP2+FLAIR	INV1+INV2
T1	731.5±9.7		106.1±3.2	55.0±3.0	126.9±3.7	23.9±3.1	30.8±3.4	32.5±3.3
MP2	837.6±11.3	106.1±3.2		161.1±4.4	233.0±4.2	130.0±4.2	136.9±4.8	73.6±3.4
T1+FLAIR	676.5±8.2	55.0±3.0	161.1±4.4		71.9±4.6	31.1±2.5	24.2±1.6	87.5±3.5
MP2+INV1	604.6±9.6	126.9±3.7	233.0±4.2	71.9±4.6		103.0±4.6	96.1±5.3	159.4±3.9
MP2+INV2	707.6±9.2	23.9±3.1	130.0±4.2	31.1±2.5	103.0±4.6		6.9±2.7	56.4±3.0
MP2+FLAIR	700.7±8.8	30.8±3.4	136.9±4.8	24.2±1.6	96.1±5.3	6.9±2.7		63.3±3.5
INV1+INV2	764.0±10.6	32.5±3.3	73.6±3.4	87.5±3.5	159.4±3.9	56.4±3.0	63.3±3.5	
WM								
T1	470.0±9.0		25.0±0.8	9.0±1.2	73.5±2.2	15.1±0.9	13.1±1.4	90.3±2.4
MP2	445.0±8.9	25.0±0.8		34.0±1.1	98.5±2.1	9.9±0.7	11.8±1.3	65.4±2.1
T1+FLAIR	479.0±9.0	9.0±1.2	34.0±1.1		64.5±2.6	24.1±1.1	22.1±0.9	99.3±2.8
MP2+INV1	543.5±10.1	73.5±2.2	98.5±2.1	64.5±2.6		88.6±2.4	86.6±2.8	163.8±2.7
MP2+INV2	455.0±8.6	15.1±0.9	9.9±0.7	24.1±1.1	88.6±2.4		1.9±1.0	75.3±2.2
MP2+FLAIR	456.9±8.6	13.1±1.4	11.8±1.3	22.1±0.9	86.6±2.8	1.9±1.0		77.2±2.7
INV1+INV2	379.7±8.0	90.3±2.4	65.4±2.1	99.3±2.8	163.8±2.7	75.3±2.2	77.2±2.7	
CSF								
T1	286.8±10.6		51.2±6.8	30.5±6.2	43.6±6.8	27.3±7.2	4.5±7.5	34.0±7.0
MP2	235.6±7.8	51.2±6.8		20.8±5.2	94.8±2.9	78.5±3.1	55.7±2.9	85.2±3.7
T1+FLAIR	256.4±9.3	30.5±6.2	20.8±5.2		74.1±5.1	57.7±5.0	34.9±4.9	64.4±5.2
MP2+INV1	330.4±8.5	43.6±6.8	94.8±2.9	74.1±5.1		16.3±3.1	39.1±3.4	9.6±3.5
MP2+INV2	314.1±9.3	27.3±7.2	78.5±3.1	57.7±5.0	16.3±3.1		22.8±2.1	6.7±2.4
MP2+FLAIR	291.3±9.1	4.5±7.5	55.7±2.9	34.9±4.9	39.1±3.4	22.8±2.1		29.5±3.0
INV1+INV2	320.8±10.0	34.0±7.0	85.2±3.7	64.4±5.2	9.6±3.5	6.7±2.4	29.5±3.0	
TIV								
T1	1488.4±22.5		29.9±7.2	76.5±4.8	9.8±6.4	11.7±5.8	39.5±6.2	23.9±5.8
MP2	1518.3±22.3	29.9±7.2		106.4±6.0	39.7±2.4	41.5±2.8	69.3±2.5	53.8±3.8
T1+FLAIR	1411.9±21.5	76.5±4.8	106.4±6.0		66.7±5.1	64.8±4.9	37.0±4.7	52.6±4.7
MP2+INV1	1478.6±21.6	9.8±6.4	39.7±2.4	66.7±5.1		1.9±1.5	29.7±1.2	14.1±2.9
MP2+INV2	1476.7±21.8	11.7±5.8	41.5±2.8	64.8±4.9	1.9±1.5		27.8±1.2	12.2±2.2
MP2+FLAIR	1448.9±21.6	39.5±6.2	69.3±2.5	37.0±4.7	29.7±1.2	27.8±1.2		15.6±2.4
INV1+INV2	1464.5±22.2	23.9±5.8	53.8±3.8	52.6±4.7	14.1±2.9	12.2±2.2	15.6±2.4	

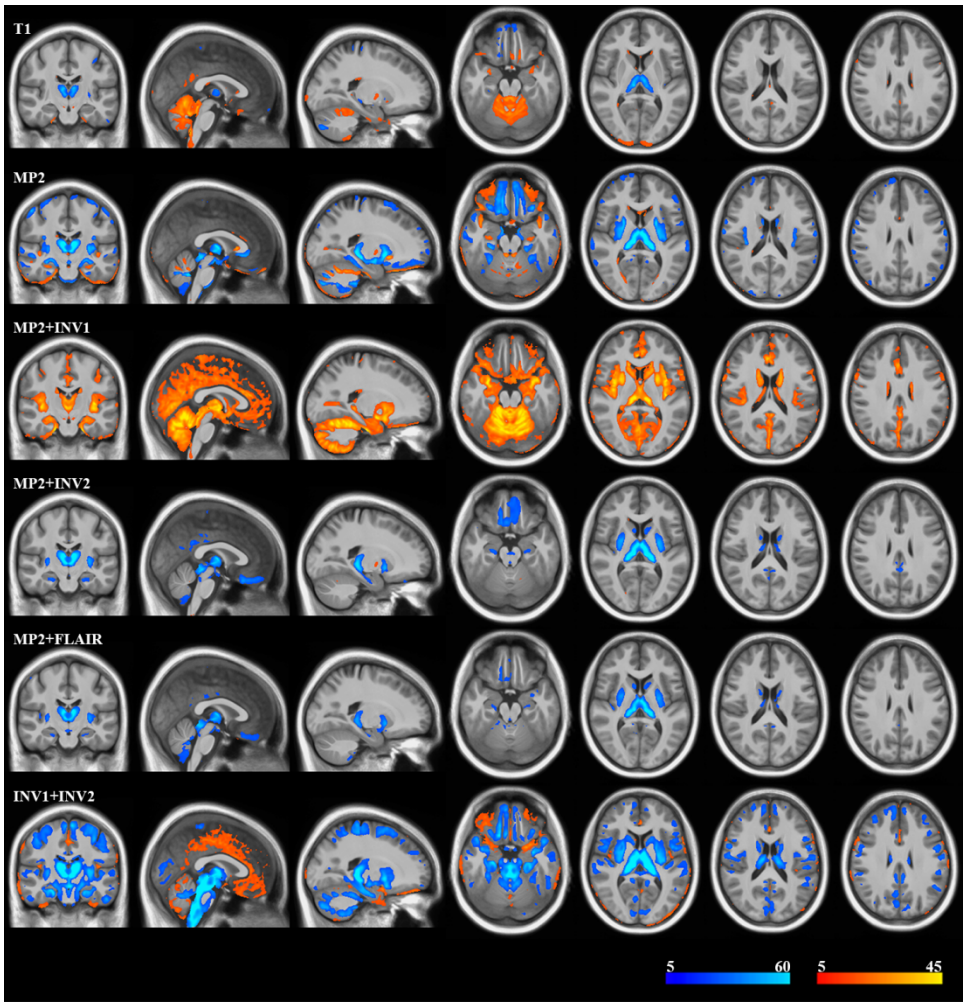
SE = standard error; shaded (shaded) are significant results, $p<0.05$.

Supplementary table 3: AUC at variable smoothing levels for all VBM models.

Model	4mm	6mm	8mm	10mm	12mm	14mm	16mm
T1	0.03	0.06	0.18	0.28	0.36	0.39	0.33
MP2	0.03	0.07	0.18	0.21	0.23	0.24	0.21
T1+FLAIR	0.03	0.05	0.17	0.25	0.33	0.36	0.32
MP2+INV1	0.03	0.05	0.19	0.44	0.45	0.38	0.29
MP2+INV2	0.03	0.09	0.13	0.22	0.25	0.31	0.32
MP2+FLAIR	0.03	0.06	0.13	0.22	0.28	0.27	0.23
INV1+INV2	0.04	0.19	0.24	0.21	0.21	0.15	0.09

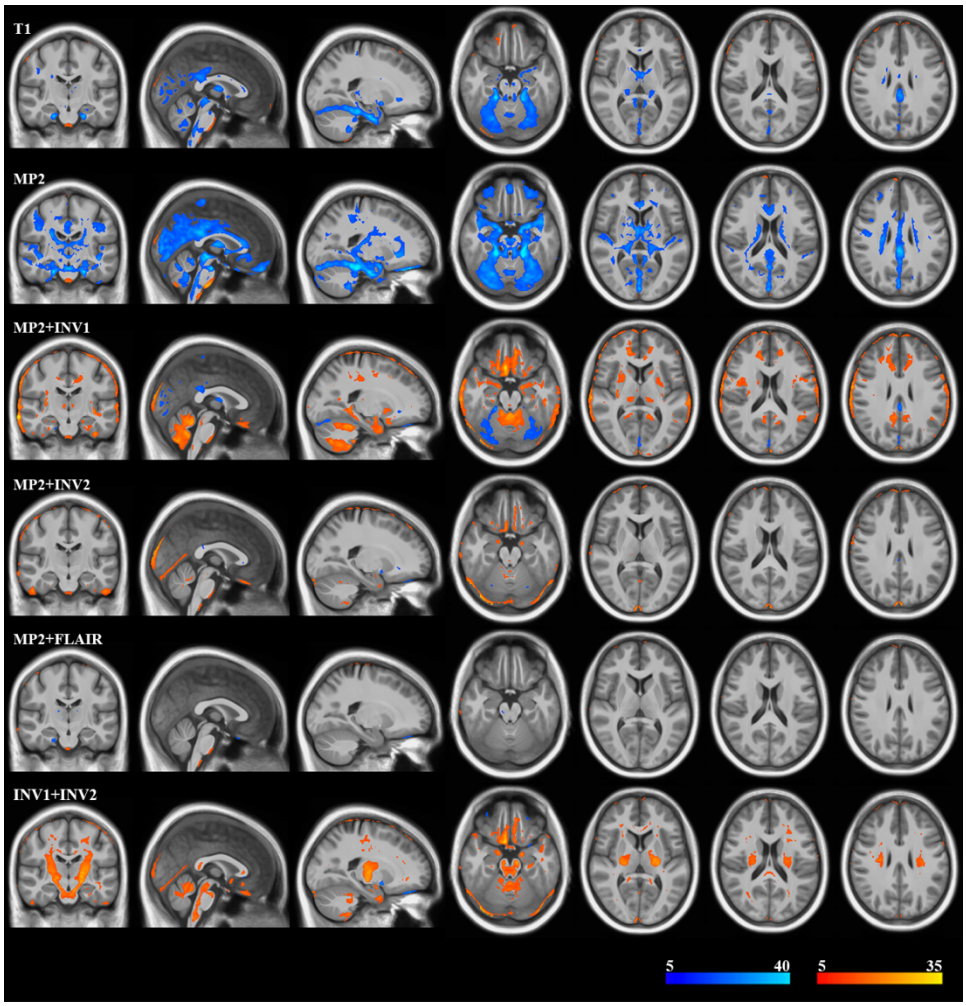
Area under curve (AUC) for all the VBM models across variable smoothing levels is present in this table. AUC for each smoothing (4mm to 16mm, step size = 2) was derived from specificity and concordant rate across all statistical cutoffs (2.5 to 6, step size = 0.1).

Supplementary figure 1: Group level differences for T1+FLAIR with rest of the models of WM increases and decreases for white matter volume (modulated images) analysis.



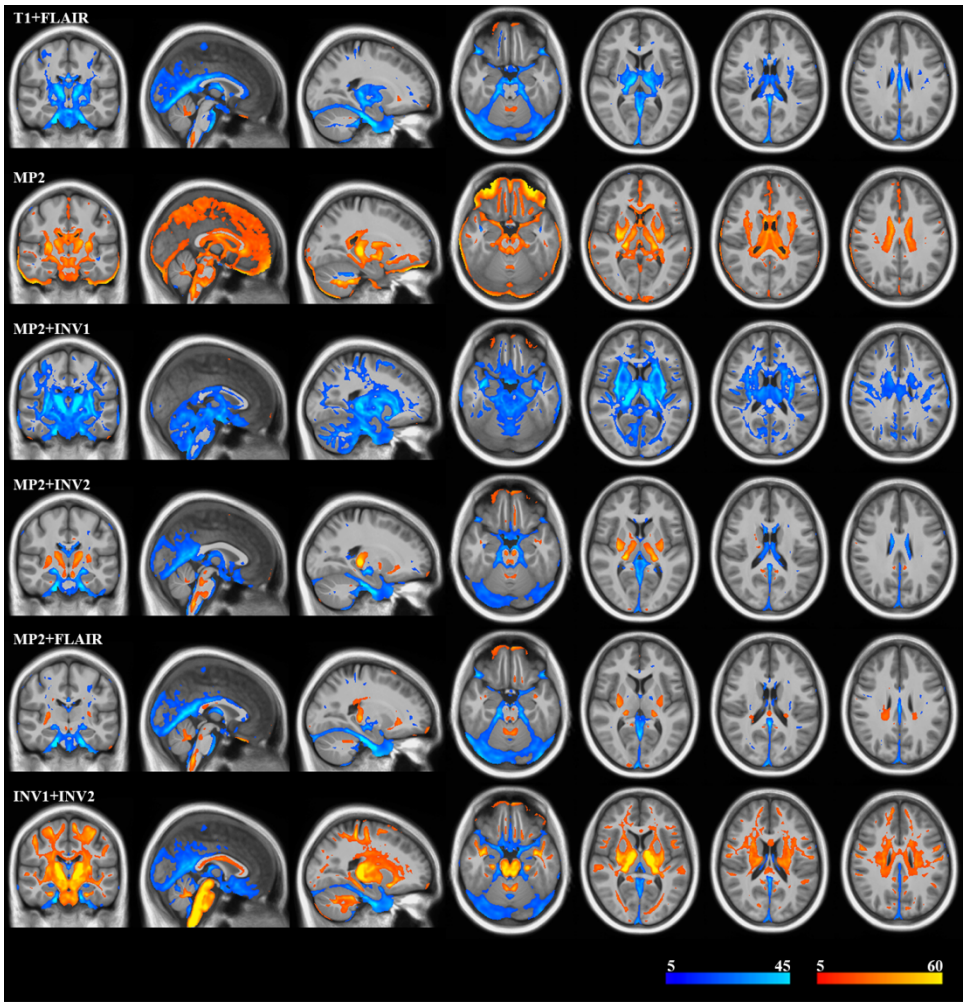
Group level comparison for T1+FLAIR with rest of the models is shown in this figure. The look-up table with Red-Yellow represents increased WM volumes in compared models against T1+FLAIR (models > T1+FLAIR), while Blue-light blue represents decreased WM volumes (models < T1+FLAIR).

Supplementary figure 2: Group level differences for T1+FLAIR with rest of the models of CSF increases and decreases for cerebrospinal fluid volume (modulated images) analysis.



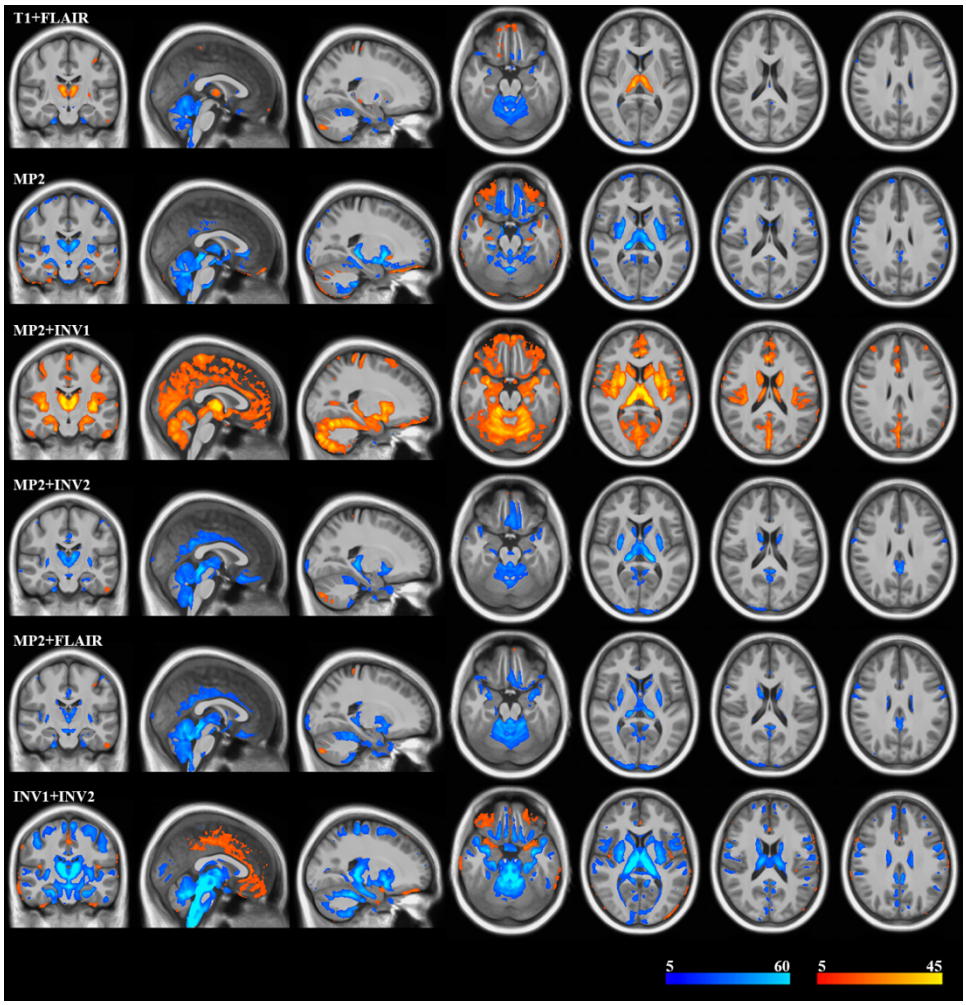
Group level comparison for T1+FLAIR with rest of the models is shown in this figure. The look-up table with Red-Yellow represents increased CSF volumes in compared models against T1+FLAIR (models > T1+FLAIR), while Blue-light blue represents decreased CSF volumes (models < T1+FLAIR).

Supplementary figure 3: Group level differences for T1 with rest of the models of GM increases and decreases for gray matter volume (modulated images) analysis.



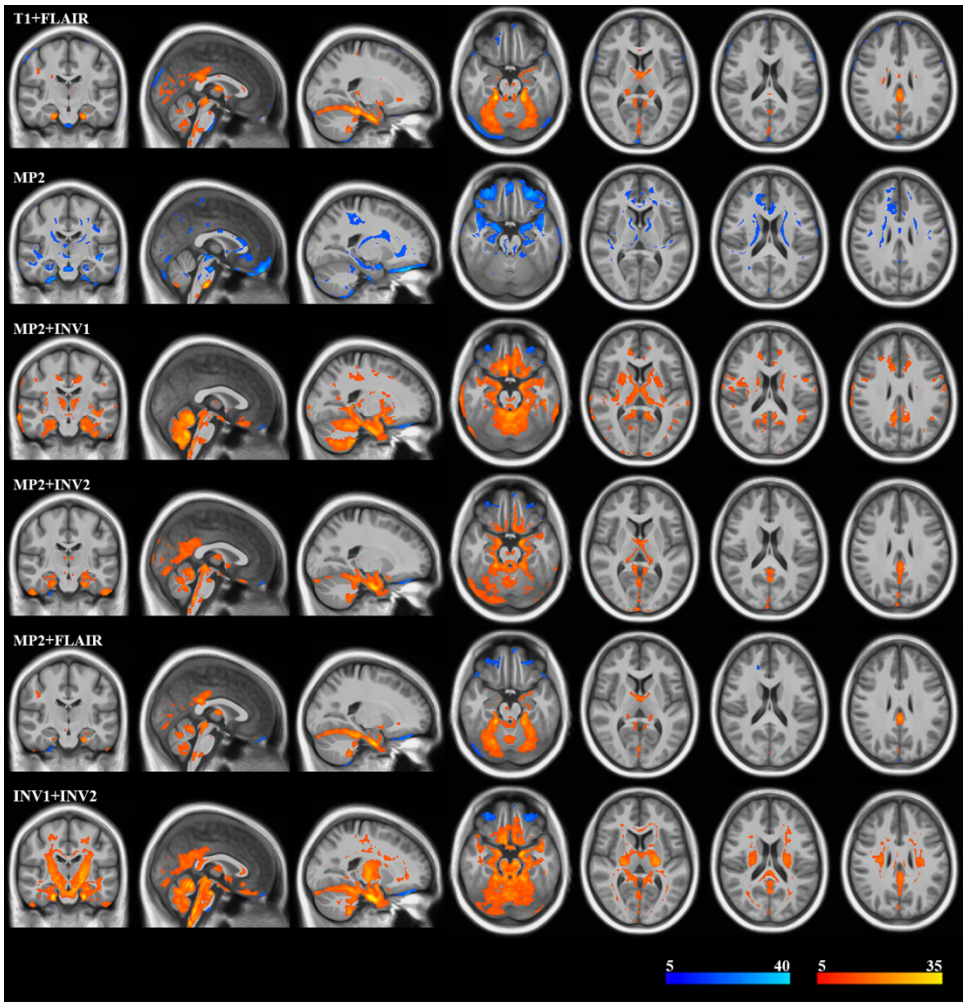
Group level comparison for T1 with rest of the models is shown in this figure. The look-up table with Red-Yellow represents increased GM volumes in compared models against T1 (models > T1), while Blue-light blue represents decreased GM volumes (models < T1).

Supplementary figure 4: Group level differences for T1 with rest of the models of WM increases and decreases for white matter volume (modulated images) analysis.



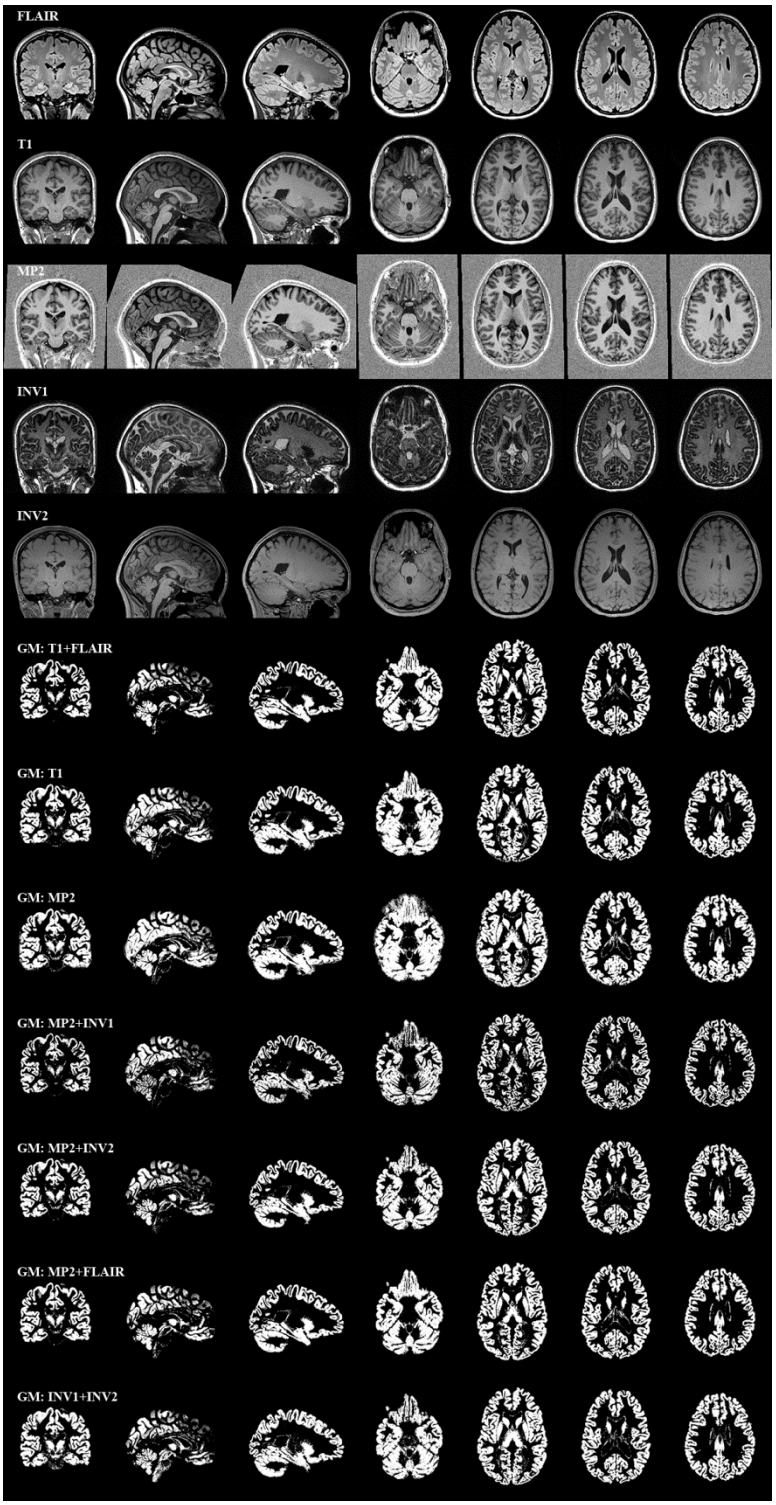
Group level comparison for T1 with rest of the models is shown in this figure. The look-up table with Red-Yellow represents increased WM volumes in compared models against T1 (models > T1), while Blue-light blue represents decreased WM volumes (models < T1).

Supplementary figure 5: Group level differences for T1 with rest of the models of CSF increases and decreases for cerebrospinal fluid volume (modulated images) analysis.



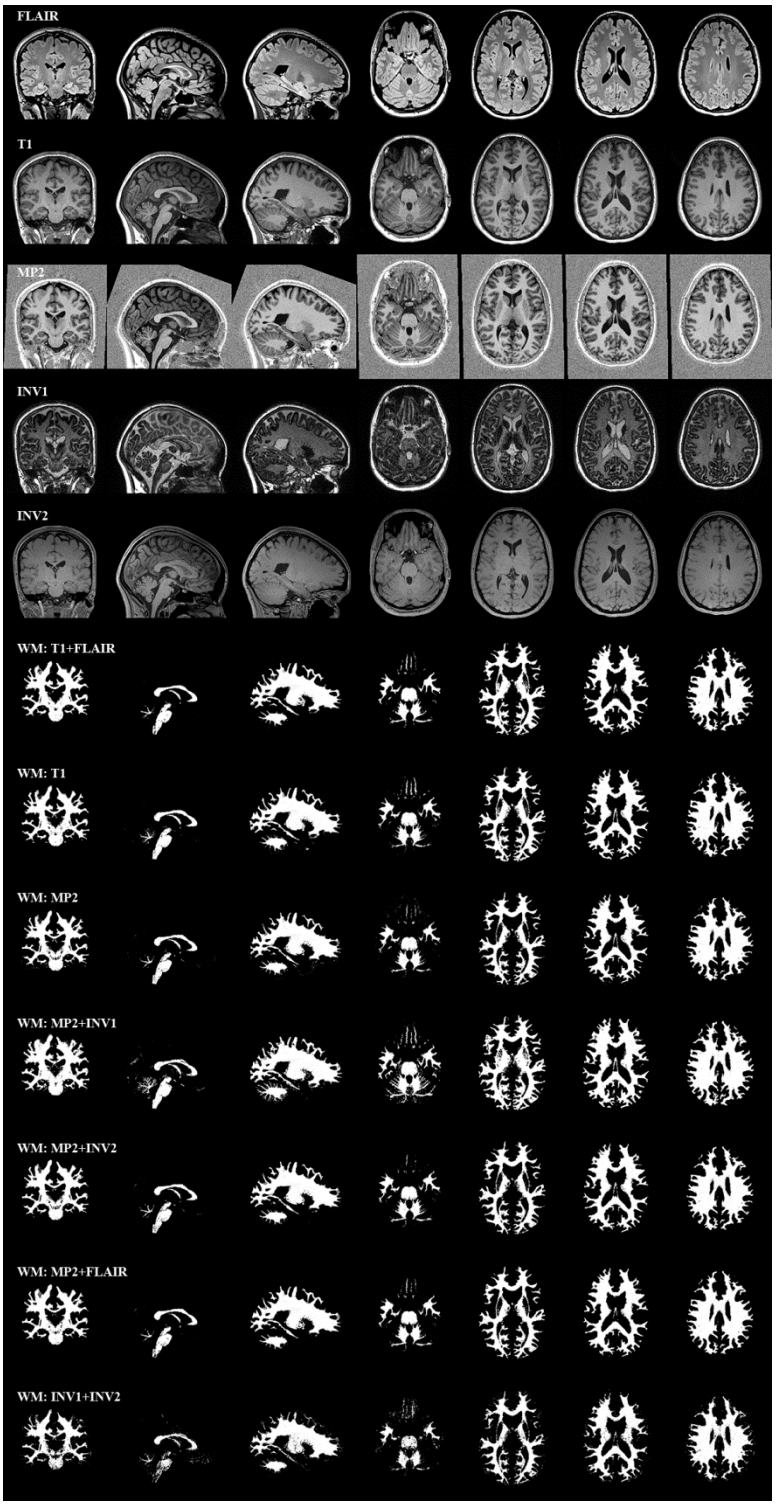
Group level comparison for T1 with rest of the models is shown in this figure. The look-up table with Red-Yellow represents increased CSF volumes in compared models against T1 (models > T1), while Blue-light blue represents decreased CSF volumes (models < T1).

Supplementary figure 6: Single-channel and multispectral segmentation results for GM in native space for all segmentation combinations.



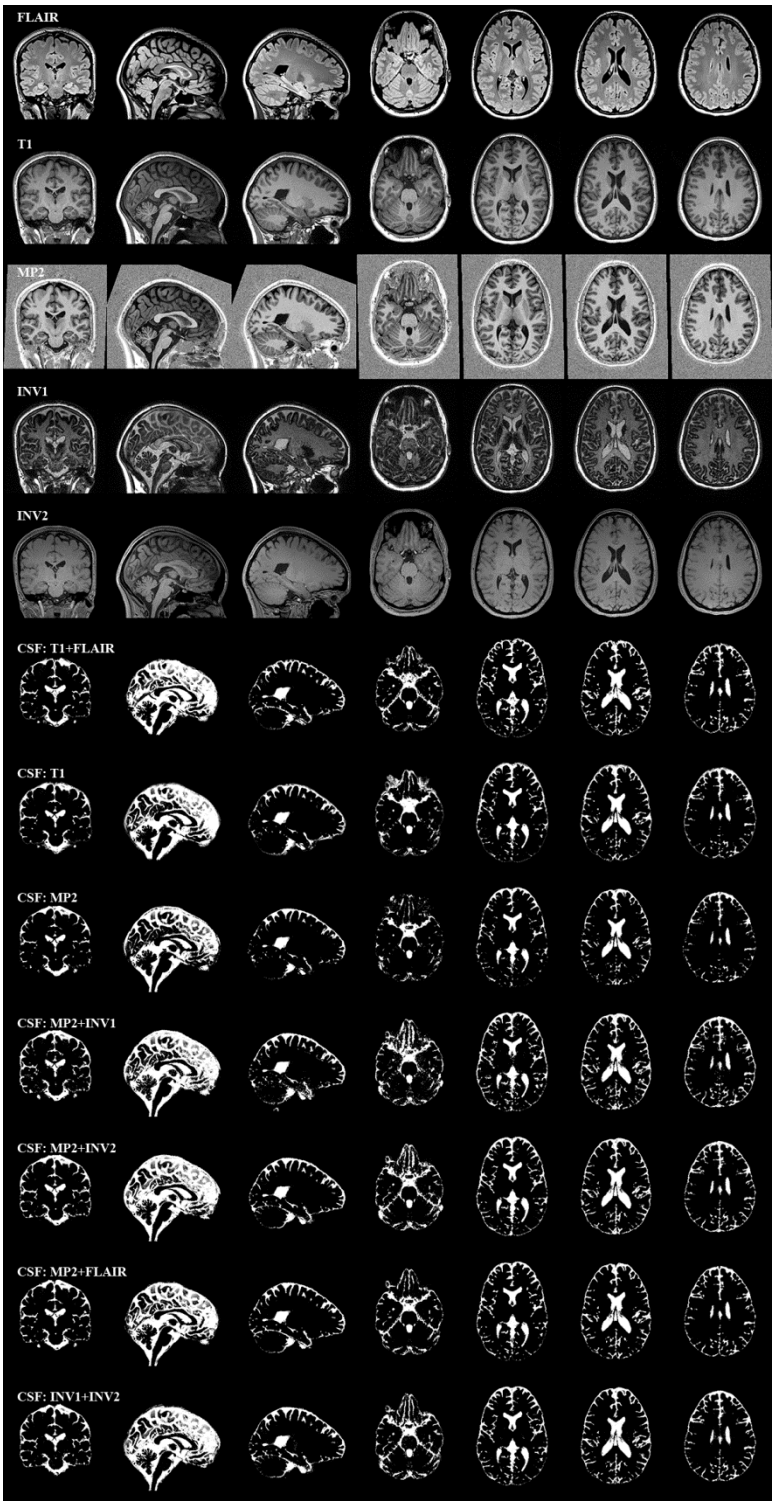
Present here are the images of gray matter segmentation for T1+FLAIR, T1, MP2, MP2+INV1, MP2+INV2, MP2+FLAIR and INV1+INV2 for a single control. Also present are the native images of FLAIR, T1, MP2, INV1 and INV2 for the same controls. These images show the translation of group results and their effect of individual subjects.

Supplementary figure 7: Single-channel and multispectral segmentation results for WM in native space for all segmentation combinations.



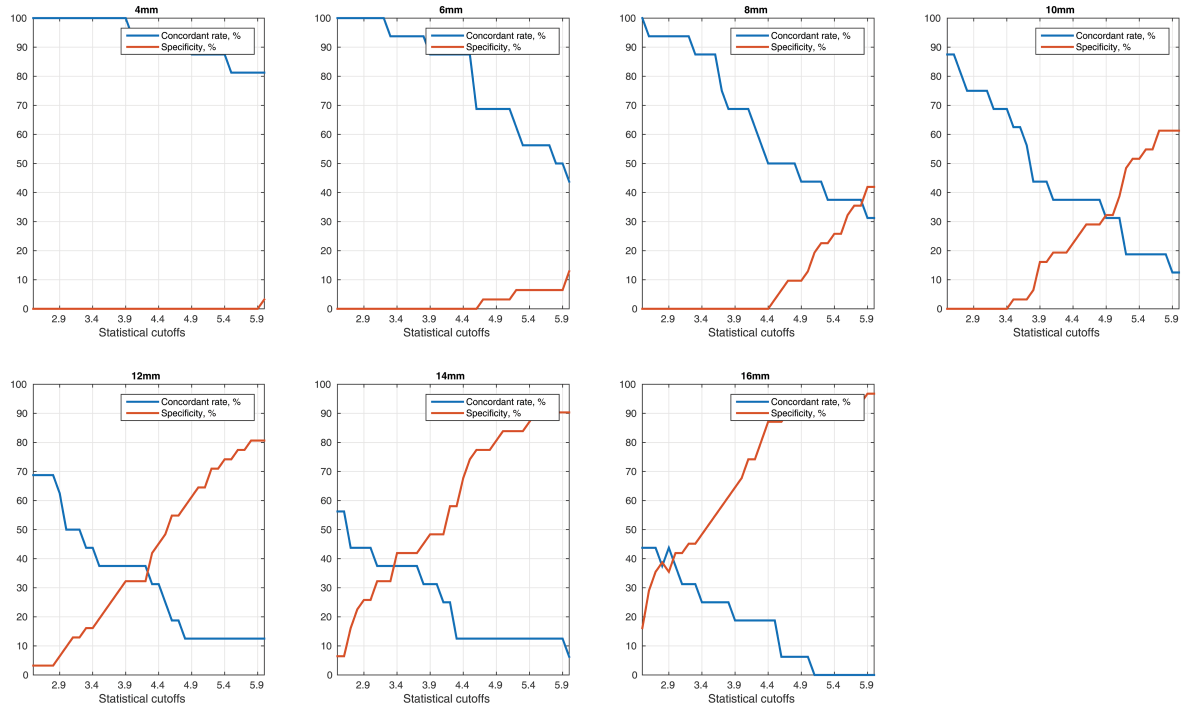
Present here are the images of white matter segmentation for T1+FLAIR, T1, MP2, MP2+INV1, MP2+INV2, MP2+FLAIR and INV1+INV2 for a single control. Also present are the native images of FLAIR, T1, MP2, INV1 and INV2 for the same controls.

Supplementary figure 8: Single-channel and multispectral segmentation results for CSF in native space for all segmentation combinations.



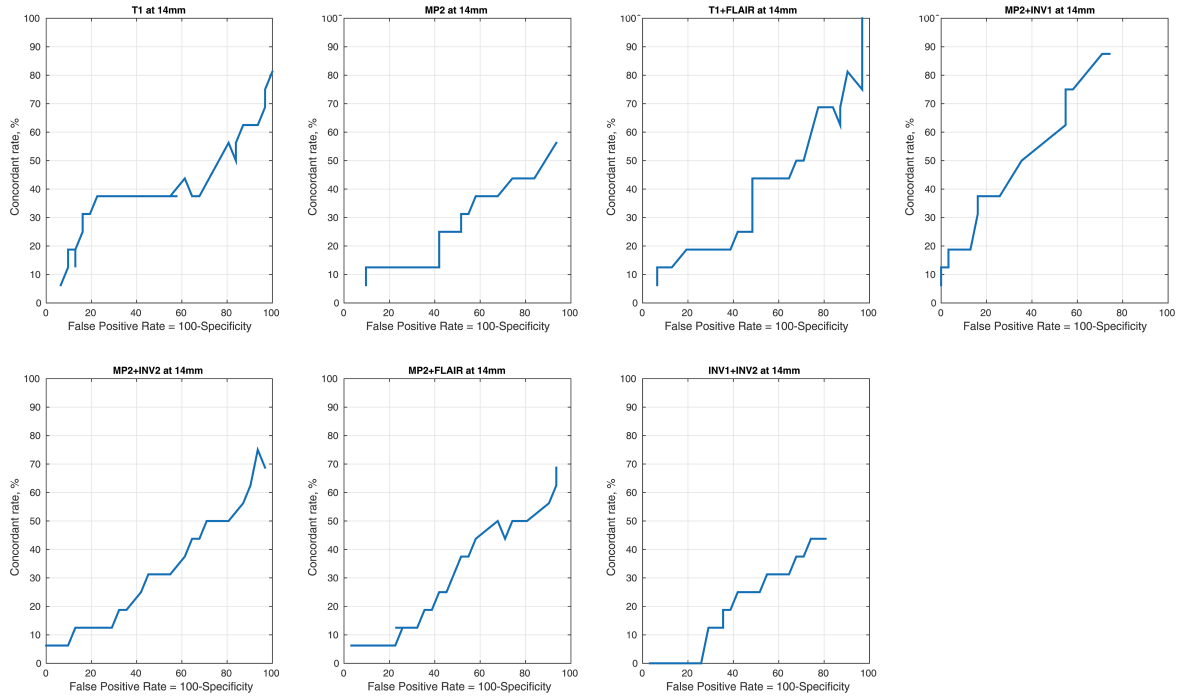
Present here are the images of cerebrospinal fluid segmentation for T1+FLAIR, T1, MP2, MP2+INV1, MP2+INV2, MP2+FLAIR and INV1+INV2 for a single control. Also present are the native images of FLAIR, T1, MP2, INV1 and INV2 for the same controls.

Supplementary figure 9: Concordant rate and specificity versus statistical cutoff plot for automated MP2 VBM findings from 4mm to 16mm smoothing kernel sizes.



Present here are the intersection plots for concordant rates and specificity versus statistical cutoffs from 2.5 to 6 and smoothing levels from 4mm to 16mm for automated MP2 VBM findings. It can be observed that as the smoothing increases, the T-threshold (intersection of concordant rate and specificity curves) shifts towards the left.

Supplementary figure 10: Receiver operating characteristic (ROC) curve for all VBM models at 14mm smoothing.



The ROC for all VBM models T1, MP2, T1+FLAIR, MP2+INV1, MP2+INV2, MP2+FLAIR and INV1+INV2 are present in this figure. The AUC calculated here represents the area occupied by the ROC curves for each VBM model.

5. Future works

The future holds infinity possibilities for the exploration and exploitation of computational methods in epilepsy. However, most importantly, there are immediate potential projects, which can be designed based on the work from this thesis. First, by taking multispectral voxel-based segmentation as a start point, several machine learning algorithms can be implemented, such as support vector computational models. Performance of multispectral guided SVMs in lesion detection can be tested in epilepsy. Second, with the technological advancements, the applicability of image analysis can be validated on ultra-high magnetic field strengths (UHF) of 9.4T. These high field strengths have an advantage over 3T, in terms of greater anatomical and functional details of the human brain / higher signal-to-noise ratio. Such information could facilitate computational methods based on this data. Third, we should aim at multimodal brain analysis where along with MRI; simultaneous information should be included from modalities such as magneto-encephalography and electro-encephalography. This would help us in elucidating epileptic findings in the brain. Last but not the least, big-data multispectral imaging analysis is a high need of the hour, where productive collaborations can be made, contributing to precision medicine in epilepsy.

6. Acknowledgment

I sincerely thank Niels Focke, for giving me the opportunity to join his group to pursue my doctoral studies. His support and contributions to my research work have been immense and I will forever be indebted to him. He has sparked in me, the spirit of competitiveness and my present and future conduct will definitely have his inspirations. I have learnt from him that, 'simply waiting is not a good strategy in science and in real life'. I thank Uwe Klose and Thomas Nägele for agreeing to be a part of my advisory committee. Their suggestions and contributions have been highly important to me. This acknowledgement is incomplete without mentioning Benjamin Bender. He has always presented himself, whenever required. Without his efforts, my doctoral studies were not a possibility. I must acknowledge that Pascal Martin has been a crucial part of my research. I sincerely thank him for his advices and suggestions offered to me, at both professional and personal levels. Bernd Kardatzaki has constantly supported me through maintain our computer clusters and keeping them up to date with the latest software for facilitating smooth and effective computational analyses. I thank him sincerely. I thank Michael Erb for his selfless contributions towards my work. I sincerely thank Petya Georgieva from the Graduate Training Centre of Tübingen. She has given me immense support and suggestions during the whole course of my doctoral studies. She has enrolled me into courses which were highly important and overlapped very well with my studies. These courses have helped me in honing my technical and interpersonal skills. I must thank Marion Kurze from University Hospital Göttingen, for bearing with my situations and facilitating my administrative requirements. I thank all the members of the Epilepsy group in Tübingen and in Göttingen. I thank both University Hospital Tübingen and University Hospital Göttingen, for doing such an incredible service to improving lives of people.

I thank each and every friend and colleague, because in one way or the other, they have been important to me. Finally, I would like to express my gratitude towards my family, for always supporting and loving me.

7. References

Aarli, J.A., 2000. Epilepsy and the immune system. Archives of neurology 57, 1689-1692.

Alarcon, G., Kissani, N., Dad, M., Elwes, R.D., Ekanayake, J., Hennessy, M.J., Koutroumanidis, M., Binnie, C.D., Polkey, C.E., 2001. Lateralizing and localizing values of ictal onset recorded on the scalp: evidence from simultaneous recordings with intracranial foramen ovale electrodes. Epilepsia 42, 1426-1437.

Alfano, B., Brunetti, A., Covelli, E.M., Quarantelli, M., Panico, M.R., Ciarmiello, A., Salvatore, M., 1997. Unsupervised, automated segmentation of the normal brain using a multispectral relaxometric magnetic resonance approach. Magnetic Resonance in Medicine 37, 84-93.

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. Neuroimage 38, 95-113.

Ashburner, J., Friston, K., 1997. Multimodal image coregistration and partitioning—a unified framework. Neuroimage 6, 209-217.

Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry--the methods. Neuroimage 11, 805-821.

Ashburner, J., Friston, K.J., 2005. Unified segmentation. Neuroimage 26, 839-851.

Barkovich, A.J., Guerrini, R., Kuzniecky, R.I., Jackson, G.D., Dobyns, W.B., 2012. A developmental and genetic classification for malformations of cortical development: update 2012. Brain 135, 1348-1369.

Beck, E., Sati, P., Sethi, V., Kober, T., Dewey, B., Bhargava, P., Nair, G., Cortese, I., Reich, D., 2018. Improved Visualization of Cortical Lesions in Multiple Sclerosis Using 7T MP2RAGE. American Journal of Neuroradiology.

Beqiri, A., Hoogduin, H., Sbrizzi, A., Hajnal, J.V., Malik, S.J., 2018. Whole-brain 3 D FLAIR at 7 T using direct signal control. Magnetic Resonance in Medicine.

Berg, A.T., Berkovic, S.F., Brodie, M.J., Buchhalter, J., Cross, J.H., van Emde Boas, W., Engel, J., French, J., Glauser, T.A., Mathern, G.W., 2010. Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005–2009. Epilepsia 51, 676-685.

Berkovic, S., Duncan, J., Barkovich, A., Cascino, G., Chiron, C., Engel, J., Gadian, D., Jackson, G., Kuzniecky, R., McLachlan, R., 1998. ILAE neuroimaging commission recommendations for neuroimaging of persons with refractory epilepsy. Epilepsia 39, 1375-1376.

Bernasconi, A., Bernasconi, N., Bernhardt, B.C., Schrader, D., 2011. Advances in MRI for'cryptogenic'epilepsies. Nature reviews neurology 7, 99.

Besson, P., Andermann, F., Dubeau, F., Bernasconi, A., 2008. Small focal cortical dysplasia lesions are located at the bottom of a deep sulcus. Brain 131, 3246-3255.

Bien, C.G., Szinay, M., Wagner, J., Clusmann, H., Becker, A.J., Urbach, H., 2009. Characteristics and surgical outcomes of patients with refractory magnetic resonance imaging–negative epilepsies. Archives of neurology 66, 1491-1499.

Blumcke, I., Thom, M., Aronica, E., Armstrong, D.D., Vinters, H.V., Palmini, A., Jacques, T.S., Avanzini, G., Barkovich, A.J., Battaglia, G., Becker, A., Cepeda, C., Cendes, F., Colombo, N., Crino, P., Cross, J.H., Delalande, O., Dubeau, F., Duncan, J., Guerrini, R., Kahane, P., Mathern, G., Najm, I., Ozkara, C., Raybaud, C., Represa, A., Roper, S.N., Salamon, N., Schulze-Bonhage, A., Tassi, L., Vezzani, A., Spreafico, R., 2011. The clinicopathologic spectrum of focal cortical dysplasias: a

consensus classification proposed by an ad hoc Task Force of the ILAE Diagnostic Methods Commission. Epilepsia 52, 158-174.
Bonilha, L., Montenegro, M.A., Rorden, C., Castellano, G., Guerreiro, M.M., Cendes, F., Li, L.M., 2006a. Voxel-based morphometry reveals excess gray matter concentration in patients with focal cortical dysplasia. Epilepsia 47, 908-915.
Bonilha, L., Montenegro, M.A., Rorden, C., Castellano, G., Guerreiro, M.M., Cendes, F., Li, L.M., 2006b. Voxel-based morphometry reveals excess gray matter concentration in patients with focal cortical dysplasia. Epilepsia 47, 908-915.
Cendes, F., Theodore, W.H., Brinkmann, B.H., Sulc, V., Cascino, G.D., 2016. Neuroimaging of epilepsy. Handbook of clinical neurology. Elsevier, pp. 985-1014.
Chin, J.H., Vora, N., 2014. The global burden of neurologic diseases. Neurology 83, 349-351.
Collins, D.L., Holmes, C.J., Peters, T.M., Evans, A.C., 1995. Automatic 3-D model-based neuroanatomical segmentation. Human brain mapping 3, 190-208.
Colliot, O., Bernasconi, N., Khalili, N., Antel, S.B., Naessens, V., Bernasconi, A., 2006. Individual voxel-based analysis of gray matter in focal cortical dysplasia. Neuroimage 29, 162-171.
Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis: I. Segmentation and surface reconstruction. Neuroimage 9, 179-194.
Delanty, N., Vaughan, C.J., French, J.A., 1998. Medical causes of seizures. The Lancet 352, 383-390.
Duncan, J.S., 1997. Imaging and epilepsy. Brain: a journal of neurology 120, 339-377.
Duncan, J.S., 2002. Neuroimaging methods to evaluate the etiology and consequences of epilepsy. Epilepsy research 50, 131-140.
Eriksson, S.H., Thom, M., Symms, M.R., Focke, N.K., Martinian, L., Sisodiya, S.M., Duncan, J.S., 2009. Cortical neuronal loss and hippocampal sclerosis are not detected by voxel-based morphometry in individual epilepsy surgery patients. Human brain mapping 30, 3351-3360.
Fiest, K.M., Sauro, K.M., Wiebe, S., Patten, S.B., Kwon, C.-S., Dykeman, J., Pringsheim, T., Lorenzetti, D.L., Jetté, N., 2016. Prevalence and incidence of epilepsy A systematic review and meta-analysis of international studies. Neurology, 10.1212/WNL. 0000000000003509.
Fisher, R.S., Boas, W.V.E., Blume, W., Elger, C., Genton, P., Lee, P., Engel Jr, J., 2005. Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). Epilepsia 46, 470-472.
Fisher, R.S., Cross, J.H., French, J.A., Higurashi, N., Hirsch, E., Jansen, F.E., Lagae, L., Moshé, S.L., Peltola, J., Roulet Perez, E., 2017. Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology. Epilepsia 58, 522-530.
Fletcher, L.M., Barsotti, J.B., Hornak, J.P., 1993. A multispectral analysis of brain tissues. Magnetic Resonance in Medicine 29, 623-630.
Focke, N.K., Bonelli, S.B., Yogarajah, M., Scott, C., Symms, M.R., Duncan, J.S., 2009. Automated normalized FLAIR imaging in MRI-negative patients with refractory focal epilepsy. Epilepsia 50, 1484-1490.
Focke, N.K., Helms, G., Kaspar, S., Diederich, C., Tóth, V., Dechent, P., Mohr, A., Paulus, W., 2011. Multi-site voxel-based morphometry—not quite there yet. Neuroimage 56, 1164-1170.
Focke, N.K., Symms, M.R., Burdett, J.L., Duncan, J.S., 2008. Voxel-based analysis of whole brain FLAIR at 3T detects focal cortical dysplasia. Epilepsia 49, 786-793.
Good, C.D., Johnsrude, I.S., Ashburner, J., Henson, R.N., Friston, K.J., Frackowiak, R.S., 2001. A voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage 14, 21-36.

Grinton, B.E., Heron, S.E., Pelekanos, J.T., Zuberi, S.M., Kivity, S., Afawi, Z., Williams, T.C., Casalaz, D.M., Yendle, S., Linder, I., 2015. Familial neonatal seizures in 36 families: clinical and genetic features correlate with outcome. Epilepsia 56, 1071-1080.

Han, J.M., Sahin, M., 2011. TSC1/TSC2 signaling in the CNS. FEBS letters 585, 973-980.

House, P.M., Lanz, M., Holst, B., Martens, T., Stodieck, S., Huppertz, H.J., 2013. Comparison of morphometric analysis based on T1- and T2-weighted MRI data for visualization of focal cortical dysplasia. Epilepsy Res 106, 403-409.

Huppertz, H.J., Grimm, C., Fauser, S., Kassubek, J., Mader, I., Hochmuth, A., Spreer, J., Schulze-Bonhage, A., 2005. Enhanced visualization of blurred gray-white matter junctions in focal cortical dysplasia by voxel-based 3D MRI analysis. Epilepsy Res 67, 35-50.

Keller, S.S., Roberts, N., 2008. Voxel-based morphometry of temporal lobe epilepsy: an introduction and review of the literature. Epilepsia 49, 741-757.

Kini, L.G., Gee, J.C., Litt, B., 2016. Computational analysis in epilepsy neuroimaging: a survey of features and methods. NeuroImage: clinical 11, 515-529.

Knake, S., Triantafyllou, C., Wald, L., Wiggins, G., Kirk, G., Larsson, P., Stufflebeam, S., Foley, M., Shiraishi, H., Dale, A., 2005. 3T phased array MRI improves the presurgical evaluation in focal epilepsies A prospective study. Neurology 65, 1026-1031.

Kotikalapudi, R., Martin, P., Marquetand, J., Lindig, T., Bender, B., Focke, N.K., 2018. (in press) Systematic assessment of multispectral voxel-based morphomtery in previously MRI-negative focal epilepsy. American Journal of Neuroradiology.

Kwan, P., Brodie, M.J., 2000. Early identification of refractory epilepsy. New England Journal of Medicine 342, 314-319.

Lambert, C., Lutti, A., Helms, G., Frackowiak, R., Ashburner, J., 2013. Multiparametric brainstem segmentation using a modified multivariate mixture of Gaussians. NeuroImage: clinical 2, 684-694.

Lindig, T., Kotikalapudi, R., Schweikardt, D., Martin, P., Bender, F., Klose, U., Ernemann, U., Focke, N.K., Bender, B., 2017. Evaluation of multimodal segmentation based on 3D T1-, T2- and FLAIR-weighted images - the difficulty of choosing. Neuroimage.

Marques, J.P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P.F., Gruetter, R., 2010. MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. Neuroimage 49, 1271-1281.

Martin, P., Bender, B., Focke, N.K., 2015. Post-processing of structural MRI for individualized diagnostics. Quant Imaging Med Surg 5, 188-203.

Martin, P., Winston, G.P., Bartlett, P., de Tisi, J., Duncan, J.S., Focke, N.K., 2017. Voxel-based magnetic resonance image postprocessing in epilepsy. Epilepsia 58, 1653-1664.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Parsons, L., Narr, K., Kabani, N., Le Goualher, G., Boomsma, D., Cannon, T., Kawashima, R., Mazoyer, B., 2001. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philos Trans R Soc Lond B Biol Sci 356, 1293-1322.

McGonigal, A., Bartolomei, F., Régis, J., Guye, M., Gavaret, M., Fonseca, A.T.-D., Dufour, H., Figarella-Branger, D., Girard, N., Péragut, J.-C., 2007. Stereoelectroencephalography in presurgical assessment of MRI-negative epilepsy. Brain 130, 3169-3183.

Mougin, O., Abdel-Fahim, R., Dineen, R., Pitiot, A., Evangelou, N., Gowland, P., 2016. Imaging gray matter with concomitant null point imaging from the phase sensitive inversion recovery sequence. Magnetic Resonance in Medicine 76, 1512-1516.

Okubo, G., Okada, T., Yamamoto, A., Kanagaki, M., Fushimi, Y., Okada, T., Murata, K., Togashi, K., 2016. MP2RAGE for deep gray matter measurement of the brain: A comparative study with MPRAGE. J Magn Reson Imaging 43, 55-62.

Pitiot, A., Totman, J., Gowland, P., 2007. Null point imaging: a joint acquisition/analysis paradigm for MR classification. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 759-766.

Pittau, F., Baud, M.O., Jorge, J., Xin, L., Grouiller, F., Iannotti, G.R., Seeck, M., Lazeyras, F., Vulliémoz, S., Vargas, M.I., 2018. MP2RAGE and Susceptibility-Weighted Imaging in Lesional Epilepsy at 7T. Journal of Neuroimaging.

Riney, C.J., Chong, W.K., Clark, C.A., Cross, J.H., 2012. Voxel based morphometry of FLAIR MRI in children with intractable focal epilepsy: implications for surgical intervention. Eur J Radiol 81, 1299-1305.

Rosenow, F., Lüders, H., 2001. Presurgical evaluation of epilepsy. Brain 124, 1683-1700.

Salmond, C., Ashburner, J., Vargha-Khadem, F., Connelly, A., Gadian, D., Friston, K., 2002. Distributional assumptions in voxel-based morphometry. Neuroimage 17, 1027-1030.

Scheffer, I.E., Berkovic, S., Capovilla, G., Connolly, M.B., French, J., Guilhoto, L., Hirsch, E., Jain, S., Mathern, G.W., Moshe, S.L., Nordli, D.R., Perucca, E., Tomson, T., Wiebe, S., Zhang, Y.H., Zuberi, S.M., 2017. ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. Epilepsia 58, 512-521.

Seiger, R., Hahn, A., Hummer, A., Kranz, G.S., Ganger, S., Küblböck, M., Kraus, C., Sladky, R., Kasper, S., Windischberger, C., 2015. Voxel-based morphometry at ultra-high fields. A comparison of 7 T and 3 T MRI data. Neuroimage 113, 207-216.

Spencer, D., 2014. MRI (minimum recommended imaging) in epilepsy. Epilepsy currents 14, 261-263.

Spencer, S.S., Williamson, P.D., Bridgers, S.L., Mattson, R.H., Cicchetti, D.V., Spencer, D.D., 1985. Reliability and accuracy of localization by scalp ictal EEG. Neurology 35, 1567-1575.

Streitbürger, D.-P., Pampel, A., Krueger, G., Lepsien, J., Schroeter, M.L., Mueller, K., Möller, H.E., 2014. Impact of image acquisition on voxel-based-morphometry investigations of age-related structural brain changes. Neuroimage 87, 170-182.

Tassi, L., Colombo, N., Garbelli, R., Francione, S., Lo Russo, G., Mai, R., Cardinale, F., Cossu, M., Ferrario, A., Galli, C., 2002. Focal cortical dysplasia: neuropathological subtypes, EEG, neuroimaging and surgical outcome. Brain 125, 1719-1732.

Taylor, D., Falconer, M., Bruton, C., Corsellis, J., 1971. Focal dysplasia of the cerebral cortex in epilepsy. Journal of Neurology, Neurosurgery & Psychiatry 34, 369-387.

Téllez-Zenteno, J.F., Hernández-Ronquillo, L., 2012. A review of the epidemiology of temporal lobe epilepsy. Epilepsy research and treatment 2012.

Tellez-Zenteno, J.F., Hernandez Ronquillo, L., Moien-Afshari, F., Wiebe, S., 2010. Surgical outcomes in lesional and non-lesional epilepsy: a systematic review and meta-analysis. Epilepsy Res 89, 310-318.

Vannier, M.W., Butterfield, R.L., Jordan, D., Murphy, W.A., Levitt, R.G., Gado, M., 1985. Multispectral analysis of magnetic resonance images. Radiology 154, 221-224.

Vezzani, A., Fujinami, R.S., White, H.S., Preux, P.-M., Blümcke, I., Sander, J.W., Löscher, W., 2016. Infections, inflammation and epilepsy. Acta neuropathologica 131, 211-234.

Visser, F., Zwanenburg, J.J., Hoogduin, J.M., Luijten, P.R., 2010. High-resolution magnetization-prepared 3D-FLAIR imaging at 7.0 Tesla. Magnetic Resonance in Medicine 64, 194-202.

Viviani, R., Pracht, E.D., Brenner, D., Beschoner, P., Stingl, J.C., Stocker, T., 2017a. Multimodal MEMPRAGE, FLAIR, and R2* Segmentation to Resolve Dura and Vessels from Cortical Gray Matter. Front Neurosci 11, 258.

Viviani, R., Stocker, T., Stingl, J.C., 2017b. Multimodal FLAIR/MPRAGE segmentation of cerebral cortex and cortical myelin. Neuroimage 152, 130-141.

Von Oertzen, J., Urbach, H., Jungbluth, S., Kurthen, M., Reuber, M., Fernandez, G., Elger, C., 2002. Standard magnetic resonance imaging is inadequate for patients with refractory focal epilepsy. Journal of Neurology, Neurosurgery & Psychiatry 73, 643-647.

Wang, Z.I., Alexopoulos, A.V., Jones, S.E., Jaisani, Z., Najm, I.M., Prayson, R.A., 2013. The pathology of magnetic-resonance-imaging-negative epilepsy. Mod Pathol 26, 1051-1058.

Wang, Z.I., Jones, S.E., Jaisani, Z., Najm, I.M., Prayson, R.A., Burgess, R.C., Krishnan, B., Ristic, A., Wong, C.H., Bingaman, W., Gonzalez-Martinez, J.A., Alexopoulos, A.V., 2015. Voxel-based morphometric magnetic resonance imaging (MRI) postprocessing in MRI-negative epilepsies. Ann Neurol 77, 1060-1075.

Wellmer, J., Quesada, C.M., Rothe, L., Elger, C.E., Bien, C.G., Urbach, H., 2013. Proposal for a magnetic resonance imaging protocol for the detection of epileptogenic lesions at early outpatient stages. Epilepsia 54, 1977-1987.

Whelan, C.D., Altmann, A., Botía, J.A., Jahanshad, N., Hibar, D.P., Absil, J., Alhusaini, S., Alvim, M.K., Auvinen, P., Bartolini, E., 2018. Structural brain abnormalities in the common epilepsies assessed in a worldwide ENIGMA study. Brain 141, 391-408.

Wiebe, S., Blume, W.T., Girvin, J.P., Eliasziw, M., 2001. A randomized, controlled trial of surgery for temporal-lobe epilepsy. New England Journal of Medicine 345, 311-318.

Wieser, H.G., 2004. Mesial temporal lobe epilepsy with hippocampal sclerosis. Epilepsia 45, 695-714.