

Electronic Thesis and Dissertation Repository

4-21-2015 12:00 AM

Multivariate Analysis of MR Images in Temporal Lobe Epilepsy

Diego H. Cantor-Rivera, *The University of Western Ontario*

Supervisor: Dr. Terry M. Peters, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree
in Biomedical Engineering

© Diego H. Cantor-Rivera 2015

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Biomedical Engineering and Bioengineering Commons](#)

Recommended Citation

Cantor-Rivera, Diego H., "Multivariate Analysis of MR Images in Temporal Lobe Epilepsy" (2015).
Electronic Thesis and Dissertation Repository. 2823.
<https://ir.lib.uwo.ca/etd/2823>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

MULTIVARIATE ANALYSIS OF MR IMAGES IN TEMPORAL LOBE
EPILEPSY

(Thesis format: Integrated Article)

by

Diego Hernando Cantor Rivera

Graduate Program in Graduate Program in Biomedical Engineering

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Diego Hernando Cantor Rivera 2015

Abstract

Epilepsy stands aside from other neurological diseases because clinical patterns of progression are unknown: The etiology of each epilepsy case is unique and so it is the individual prognosis. Temporal lobe epilepsy (TLE) is the most frequent type of focal epilepsy and the surgical excision of the hippocampus and the surrounding tissue is an accepted treatment in refractory cases, especially when seizures become frequent increasingly affecting the performance of daily tasks and significantly decreasing the quality of life of the patient. The sensitivity of clinical imaging is poor for patients with no hippocampal involvement and invasive procedures such as the Wada test and intracranial EEG are required to detect and lateralize epileptogenic tissue.

This thesis develops image processing techniques using quantitative relaxometry and diffusion tensor imaging aiming to complement current MRI-based diagnosis by providing additional diagnostic clues. Chapter 2 develops the concept of individual feature maps on regions of interest. A laterality score on these maps correctly distinguished left TLE from right TLE in 12 out of 15 patients. Chapter 3 explores machine learning models to detect TLE, obtaining perfect classification for patients with left TLE, and 88.9% accuracy for patients with right TLE. Chapter 4 focuses on temporal lobe asymmetry developing a voxel-based method for assessing asymmetry and verifying its applicability to individual predictions (92% accuracy) and group-wise statistical analyses. Informative ROI and voxel-based informative features are described for each experiment, demonstrating the relative importance of mean diffusivity over other MR imaging alternatives in identification and lateralization of TLE patients. Finally, the conclusion chapter discusses contributions, main limitations and outlining options for future research.

Keywords: TLE, Q-MRI, DTI, FA, MD, T1, T2, DESPOT1, DESPOT2, asymmetry, Kolmogorov-Smirnov test, single-subject t-test, epilepsy, SVM, machine learning, feature selection, support vector machines, diffusion tensor imaging, quantitative relaxometry, quantitative imaging

A mi papá, Luis Hernando Cantor Casas

Acknowledgements

First and foremost I would like to thank Dr. Terry Peters for having welcomed me to be a part of his team. Throughout these years he has been a mentor and has taught me to be patient, dedicated and to think outside the box. He has infinite patience and he is always been available for his students, even when away on conferences. Terry allowed me to take time off to write a book, even when that meant losing a student for a year. He recognized the importance of this project and its significance in my professional career, and I owe it to him. I can proudly say that I am a better professional and a better person thanks to him.

I would also like to offer my special thanks to Dr. Ali Khan. Ali came along at a moment when the epilepsy project was hitting a logistical end of the road, and he single-handedly turned it around and made it into the project that it is today in which effective clinical collaboration occurs and students have access to clinicians. This had a tremendous impact on our group and directly on my research. Ali has taught me to be critical of my own work and to not lock myself into a single perspective. He is true to the spirit of science, and I admire that in him. He has been directly involved on the day to day of my project and his structured approach and oversight have resulted in great work quality in which I take pride.

I wish to acknowledge the help provided by Dr. Sandrine de Ribaupierre, Dr. Seyed Mirsattari, and Dr. Robert Bartha in my advisory committee. Thank you for reviewing my manuscripts, for evaluating my knowledge on the different epilepsy topics, and for giving me your professional insight and guidance. Similarly, I want to thank Dr. Jonathan Lau, for his help, his comments on my manuscripts and for his friendship. Similarly, I want to thank Dr. Jorge Burneo and Dr. David Steven for participating in our 7 a.m. Friday meetings! and for their support and contributions to our epilepsy research project.

I am particularly grateful for the assistance given by Diana Timmermans and Christine Ellwood from the Biomedical Engineering program at Western, and by Anna Pimentel, Sherri Couto and Janette Wallace at Robarts. These people have helped me tremendously during all these years with innumerable administrative tasks: guys, you go the extra-mile for the students.

To my friends, Danielle Pace and Chris Russ, I profoundly admire your dedication, intel-

lect, and sense of humour. I am honoured to be your friend. Cheers!

Special thanks to John Baxter, I really, really have enjoyed having you as a lab-neighbour. All our conversations have been tremendously enriching for me. I have learned a lot from you, you inspire me and challenge me to be a better engineer. Thank you very much for taking the time to review my manuscripts and being always willing to have academic discussions. Anyone who has the opportunity of working with you is very lucky and I am fortunate to be one of them.

Golafsoun Ameri, thank you so much for being my friend. You have given me strength throughout all this process. I admire you, you are a great engineer and a fantastic person.

To all my friends and labmates at Robarts. Sadly I cannot mention you all but you know who you are. Thank you for sharing the ride and for being a lighthouse when the sea is stormy.

Marcela Hernandez-Hoyos, Marcela, I would not be here without you. You have had a tremendous impact in my life and I am grateful to have had you as a mentor and now as a colleague.

To my mentors who have enriched my academic life and have taught me so much: Maciej Orkisz and Oscar Acosta in France, Olivier Salvado in Australia and Sergio Rojas, Rubby Casallas, and Jorge Villalobos in Colombia. I am grateful for having had the opportunity to work and learn from you.

Finally I want to thank my family for their patience and support. Jose my partner, thank you so much for your being there for me every step of the way. There are no words in English or Spanish to describe my gratitude for your infinite love, patience and support. It does not matter that the night is cold as long as I have you by my side. To my mom Cecy, and my brothers Jonathan and Fredy, you are the reason. And lastly to my dad, Hernando Cantor, I know you would have been so immensely proud of me. I miss you.

Contents

Abstract	ii
Acknowledgements	iv
List of Figures	xii
List of Tables	xv
List of Appendices	xvi
List of Abbreviations, Symbols, and Nomenclature	xvii
1 An introduction to the analysis of temporal lobe epilepsy with magnetic resonance imaging	1
1.1 Temporal Lobe Epilepsy: clinical overview	2
1.1.1 Seizures	4
Localization of temporal lobe seizures	5
1.1.2 Role of Electroencephalography	6
Intracranial EEG	8
1.1.3 Role of Imaging	9
Structural MRI	9
Functional MRI	11
PET, SPECT and MRS	12
1.1.4 Surgery: clinical challenges and current techniques	14
1.2 MRI research in Temporal Lobe Epilepsy	15

1.2.1	Quantitative Relaxometry	15
	T1- and T2-weighted images	15
	Obtaining quantitative T1 and T2 maps	16
	Accelerated methods for T1 and T2 measurement	17
1.2.2	Diffusion tensor imaging	19
	Imaging principle: diffusion weighted imaging	19
	The diffusion tensor	20
	Technical considerations	21
	Mean Diffusivity and Fractional Anisotropy	21
	MD and FA in the study of TLE	22
1.3	Imaging Analysis Techniques	23
1.3.1	Region-of-interest techniques (ROI)	23
1.3.2	Voxel-based statistical techniques	23
	Voxel-based morphometry (VBM)	24
	Voxel-based relaxometry (VBR)	26
	Voxel-based statistics of DTI-derived maps (VBS)	28
1.3.3	Structural analysis of white matter	29
	Tract-based spatial statistics (TBSS)	29
	White matter tractography and tractometry	30
1.4	Machine Learning	32
1.4.1	Feature selection	33
	Curse of dimensionality	33
	Supervised feature selection	34
	Unsupervised feature selection	35
1.4.2	Support Vector Machines	36
	Principle	36
	Kernels and the kernel trick	38
	Soft margins	39

Training and testing strategies	40
Measurements of classification performance	41
1.4.3 Support vector machines in neuroimaging studies of TLE	42
Multivariate pattern analysis	43
1.5 Thesis objectives and scope	44
1.5.1 Thesis overview	44
Chapter 2: Characterization of TLE through individual patient studies .	45
Chapter 3: TLE detection using support vector machines in multi- parametric quantitative MR imaging	45
Chapter 4: Classification of temporal lobe epilepsy patients using voxel- based asymmetry	46
Chapter 5: Conclusions	46
1.6 Co-authorship Statement	47
2 Characterization of TLE through individual patient studies	61
2.1 Introduction	61
2.2 Method	62
2.2.1 Participants and Imaging Protocol	62
2.2.2 Image Preprocessing	64
2.2.3 Segmentation of regions of interest (ROI)	64
2.2.4 Feature quantification	65
Measuring asymmetry in ROI intensities	66
2.2.5 Single-subject statistical tests	69
2.2.6 Individual Feature Maps	70
2.2.7 Patient lateralization based on temporal lobe asymmetry	71
2.2.8 Group analysis	71
2.3 Results	72
2.3.1 Individual feature maps for a MTS patient	74
2.3.2 Individual feature maps for a non-MTS patient	74

2.3.3	Group analysis	77
	Ia - Intensity abnormalities across regions of interest	77
	Ib - Intensity abnormalities by image type	78
	IIa - Volume asymmetry by region of interest	79
	IIb - Ipsilateral/contralateral volume reduction	81
	Group-wise comparison of hippocampal volumes	81
2.4	Discussion	82
2.4.1	Study design considerations and limitations of this work	82
2.4.2	On the role of single-subject statistics	83
2.4.3	Comparison with similar studies	84
2.4.4	Clinical Relevance	85
2.5	Conclusion	86
3	TLE detection using support vector machines in multi-parametric quantitative	
	MR imaging	93
3.1	Introduction	93
3.2	Methods	94
3.2.1	Overview	94
3.2.2	Participants	96
3.2.3	Imaging Protocol and Preprocessing	96
3.2.4	Feature Extraction	97
3.2.5	Support Vector Machines	98
3.2.6	Experiments	99
3.2.7	Classification Models	100
3.2.8	Feature Selection	101
	Training Set Subsampling	101
	Correlation-Based Feature Selection	102
	ANOVA-Based Feature Selection	104
	Voting Strategy	105

3.2.9	Dimensionality reduction	105
3.2.10	SVM selection and parameter optimization	106
3.3	Results	107
3.3.1	Detection of TLE patients	108
3.3.2	Detection of L-TLE patients	112
3.3.3	Detection of R-TLE patients	113
3.3.4	Evaluation of feature selection methods	115
3.3.5	Informative features for TLE detection	117
3.3.6	Effect of the sample size on the reliability of the results	118
3.4	Discussion	119
3.4.1	Analysis of informative features	119
3.4.2	Criteria for the selection of the feature set cardinality	121
3.4.3	Criteria for the selection of a classification model	123
3.4.4	Comparison with similar studies	124
3.4.5	Limitations of this work	125
3.5	Conclusion	126
4	Classification of temporal lobe epilepsy patients using voxel-based asymmetry	132
4.1	Introduction	132
4.2	Methods and Materials	133
4.2.1	Participants	133
4.2.2	Data acquisition and preprocessing	134
4.2.3	Measuring Voxel-based asymmetry	136
Defining homologous voxel neighbourhoods	136	
Measuring voxel-based asymmetry	137	
4.2.4	Statistical analysis of temporal lobe asymmetry	140
4.2.5	TLE patient identification using Support Vector Machines	141
Feature selection	142	
Robust features and informative features	143	

4.2.6	Baseline experiment	144
4.3	Results	144
4.3.1	Differences in temporal lobe asymmetry between patients and con trols	144
4.3.2	TLE patient identification using SVM on voxel-based asymmetry features	145
4.3.3	Analysis of informative features	146
	Informative features in the mesial temporal region (MSL)	149
	Informative features in the lateral temporal region (NEO)	150
	Informative features in the anterior temporal lobe (ATL)	151
4.4	Discussion	152
4.4.1	Comparison to multi-voxel pattern analysis methods	153
4.4.2	Comparison to similar multi-parametric classification studies	153
4.4.3	Feature selection trade-offs	154
4.4.4	Criteria for the selection of K	155
4.4.5	Limitations	156
4.5	Conclusion	157
5	Conclusions	163
5.1	Contributions	163
5.2	Limitations	165
5.3	Future Research	166
A	Individual Feature Maps	171
A.1	IFM for a MTS patient	171
A.2	IFM for a non-MTS patient	175
B	Research Ethics Board Approval	179
	Curriculum Vitae	181

List of Figures

1.1	Anatomy of the temporal lobe	2
1.2	Neurotransmission	4
1.3	Hypothesis of multiple independent generators in TLE	5
1.4	EEG principle	7
1.5	10-20 electrode array	8
1.6	Temporal lobe seizure	9
1.7	Structural imaging in TLE	10
1.8	PET-FDG	13
1.9	Temporal lobe surgery	14
1.10	Quantitative Relaxometry	18
1.11	Measuring diffusion	20
1.12	MD and FA	22
1.13	Voxel-based statistical techniques	24
1.14	VBM and VBR comparison	27
1.15	Tractography	30
1.16	Supervised feature selection	35
1.17	Support Vector Machines	37
2.1	Kolmogorov-Smirnov test	66
2.2	Intensity distribution asymmetry	67
2.3	Measuring intensity asymmetry	68
2.4	Statistical tests	69
2.5	IFM for a L-TLE patient with MTS	75

2.6	IFMs for a R-TLE patient without MTS	76
2.7	Intensity abnormalities by group	78
2.8	Intensity abnormalities by image type	79
2.9	Comparison of volume asymmetry abnormalities between groups	80
2.10	Volume asymmetry (contralateral vs. ipsilateral)	81
3.1	Overview of the method	95
3.2	Image preprocessing and feature extraction steps	98
3.3	Linear vs. non-linear SVMs	99
3.4	Feature selection	102
3.5	Grid search for RBF SVM parameter selection	107
3.6	Parameter selection	108
3.7	Stable ROI selected based on left/right mean intensity	110
3.8	Stable ROI selected based on asymmetry	111
3.9	Feature space projection	111
3.10	Feature stability	116
3.11	Similarity between feature selection methods	117
3.12	SVM reliability index	119
3.13	Relevant features by kind	120
3.14	Classification accuracy vs. number of features	122
3.15	Computational performance	124
4.1	Method overview	134
4.2	Establishing hemispheric correspondence	136
4.3	Evaluation of voxel-based asymmetry	138
4.4	Construction of voxel-based asymmetry maps (VBA maps)	139
4.5	Visual inspection of asymmetry maps	140
4.6	Temporal lobe regions	141
4.7	Feature selection	143

4.8 ROI asymmetry differences between TLE patients and controls 146

4.9 SVM classification 147

4.10 Classification accuracy as a function of feature set cardinality 149

4.11 Analysis of informative temporal lobe features 151

5.1 Subject-to-boundary correlation 167

5.2 voxel-based asymmetry patterns 168

List of Tables

1.1	Examples of kernel functions	39
2.1	Clinical summary	63
2.2	IFM results	73
3.1	TLE detection results	109
3.2	L-TLE detection results	112
3.3	R-TLE detection results	114
3.4	Summary of ROI identified by informative features	121
3.5	Estimation of K using a L1-penalized logistic regression model	122
4.1	Clinical summary and patient demographics	135
4.2	Descriptive statistics	144
4.3	Multivariate test results	145
4.4	Patient identification using SVMs	148
4.5	Analysis of informative features	150

List of Appendices

Appendix A Individual Feature Maps	171
Appendix B Research Ethics Board Approval	179

List of Abbreviations, Symbols, and Nomenclature

AED antiepileptic drug.

CSF cerebrospinal fluid.

ECDF empirical cumulative distribution function.

EEG electroencephalography.

FA fractional anisotropy.

FCD focal cortical dysplasia.

FLAIR fluid attenuation inversion recovery.

GM gray matter.

ICA independent component analysis.

ILAE International League Against Epilepsy.

KS Kolmogorov-Smirnov test.

L-TLE left TLE.

MD mean diffusivity.

MRS magnetic resonance spectroscopy.

MTL mesial temporal lobe.

MTS mesial temporal sclerosis.

MTS+ group of patients with mesial temporal sclerosis.

MTS- group of patients without mesial temporal sclerosis.

MVPA multi-voxel pattern analysis.

PCA principal component analysis.

PET positron emission tomography.

R-TLE right TLE.

SPECT single-photon emission computed tomography.

SPGR spoiled gradient recalled echo.

SVM support vector machine.

TLE temporal lobe epilepsy.

VBA voxel-based asymmetry.

VBM voxel-based morphometry.

VBR voxel-based relaxometry.

WM white matter.

Chapter 1

An introduction to the analysis of temporal lobe epilepsy with magnetic resonance imaging

Overview This chapter provides an introduction to TLE with a clinical overview where electroencephalography and neuroimaging are introduced as the main diagnostic tools. This offers the context for the discussion of current MRI research in TLE, motivating the use of the two techniques in this work: quantitative relaxometry and diffusion tensor imaging. Subsequently, a brief description of image analysis, and machine learning techniques relevant to this thesis are included. The chapter concludes with a description of the goals and scope of this work and a summary of each chapter.

1.1 Temporal Lobe Epilepsy: clinical overview

Temporal lobe epilepsy (TLE) is the most common type of focal epilepsy, accounting for up to 60% of all adult epilepsy cases [1]. TLE is characterized by seizures originated in the temporal lobe (Figure 1.1).

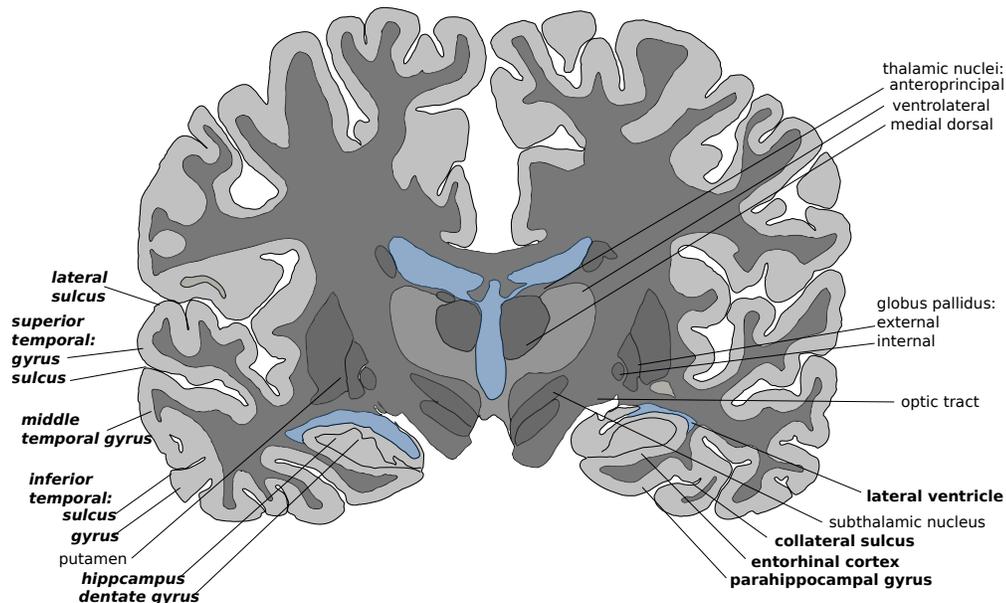


Figure 1.1: **Anatomy of the temporal lobe.** This schematic shows temporal lobe structures in bold, in relationship with limbic and extra-temporal structures. ¹

According to the International League Against Epilepsy (ILAE) having a seizure is not a sufficient condition for an epilepsy diagnosis. The seizures must be “*frequent, and unprovoked by any immediate identified cause*” [1]. According to the ILAE classification of seizures from 1981, focal seizures can be: partial simple (no loss of consciousness), or partial complex (loss of consciousness) [2]. However, this classification has been debated by experts around the world, due to the lack of agreement in the embodiment of the word consciousness. After all, what does it mean to be conscious? With this philosophical dilemma at hand, the ILAE task force shifted the emphasis of the classification from consciousness to *cognition*, encompassing five categories [3]:

¹Adapted with permission from <http://www.brains.rad.msu.edu>, and <http://brainmuseum.org>, supported by the US National Science Foundation.

- perception: the symbolic conception of sensory information
- attention: the appropriate selection of a principal perception or task
- emotion: the appropriate affective significance of a perception
- memory: the ability to store and retrieve percepts or concepts, and
- executive function: the process of anticipation, selection, monitoring of consequences, and initiation of motor activity including praxis, and speech.

With this definition, the ambiguity of the term *loss of consciousness* could be replaced with the more practical *loss of cognitive ability* or, in other words, partial complex seizures could be replaced by partial *dyscognitive* seizures, as long as two or more cognitive categories were deemed impaired during the seizure [4]. Nonetheless, this classification of partial seizures has been heavily criticized [5], as there is no consensus that *cognition* is more inclusive or even more useful than *consciousness*. Under these circumstances, the ILAE's stance is that of continuity with the original 1981 classification, suggesting an additional characterization of focal seizures on a case-by-case basis [5, 6].

The multiplicity of taxonomies and the lack of a consistent ontology are not caused only by scientific stances or clinical disagreements. There is an undeniable reality: temporal lobe epilepsy is indeed a *heterogeneous disease*, with multiple etiologies and abnormal physiological processes that we do not fully understand. Electroencephalography (EEG) and neuroimaging are the main tools that have allowed us to better understand TLE, and among imaging modalities, structural MRI is the most commonly used. Both EEG and MRI have sensitivity and specificity issues and patient evaluation is performed using both: agreement between EEG and MRI translates into optimal clinical outcomes; disagreement, into long-term or *invasive* EEG monitoring, and palliative care. Thus, the goal of any TLE diagnostic technique is to detect and pinpoint the origin of seizures. But, what are seizures?

1.1.1 Seizures

A *seizure* (from the Latin *sacire* -to take possession of) is the clinical manifestation of an abnormal, excessive, hyper-synchronous discharge of a population of cortical neurons [7]. Such discharges are caused by either excessive excitation or loss of inhibition in neuronal synapses (Figure 1.2).

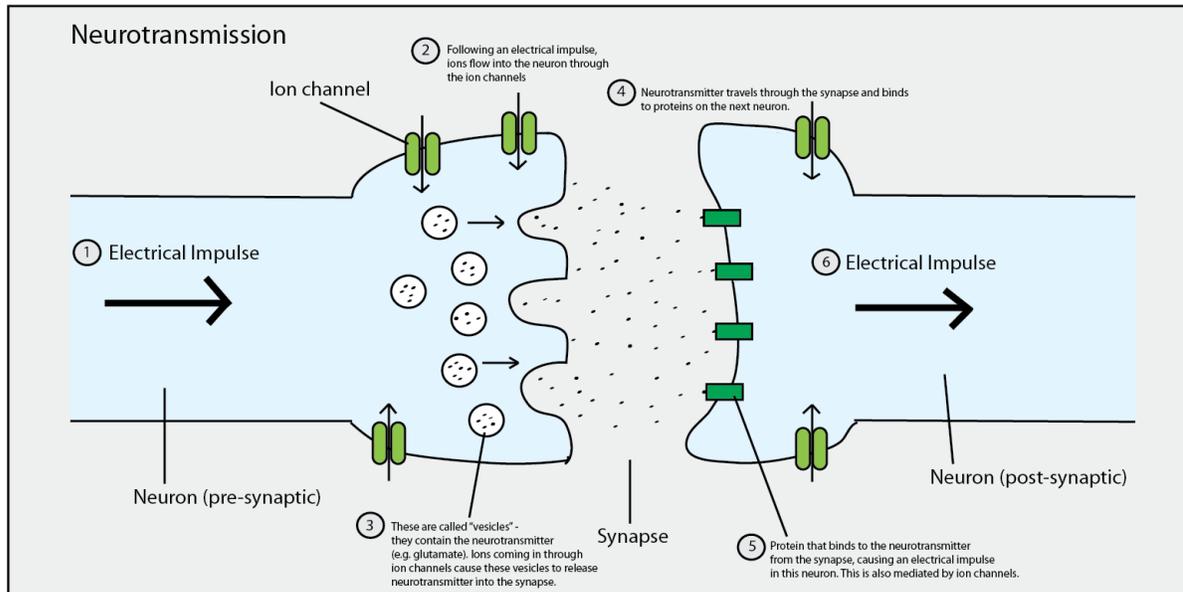


Figure 1.2: **Neurotransmission.** Neurotransmission occurs by the interaction of neurotransmitter chemicals with ion channels in the neuronal synapses ².

When the action potential (electrical impulse) reaches the axon terminal, voltage-gated ion channels open, allowing the influx of Na^+ ions. These ions cause the release of chemicals known as neurotransmitters into the synapses. Neurotransmitters bind to receptors present in the post-synaptic neuron, causing a change of permeability and opening (excitatory action) or closing (inhibitory action) the post-synaptic neuron ion channels (Na^+ , K^+ , Cl^- , Ca^{2+}). A new action potential can be induced or inhibited in the post-synaptic neuron depending on the type and the relative concentration of ions between the pre-synaptic and the post-synaptic neuron. After accomplishing their task, neurotransmitters are broken down by enzymes in the synaptic cleft. The induced action potential is “all-or-none”: the post-synaptic neuron must reach a

²Diagram courtesy of Emily Coyte, twitter: @emilycoyte

threshold in order to fire and propagate the message.

In this context, seizures are associated with malfunction of the ion channels modulated by the excess (or the lack of) neurotransmitters, in particular: glutamate (excitatory neurotransmitter) and GABA (inhibitory neurotransmitter). It has been shown in animal models that glutamate agonists (substance capable of activating glutamate receptors) can induce seizures [8, 9], whereas GABA agonists such as benzodiazepines and barbiturates are the base for antiepileptic drugs (AEDs) [10].

Localization of temporal lobe seizures

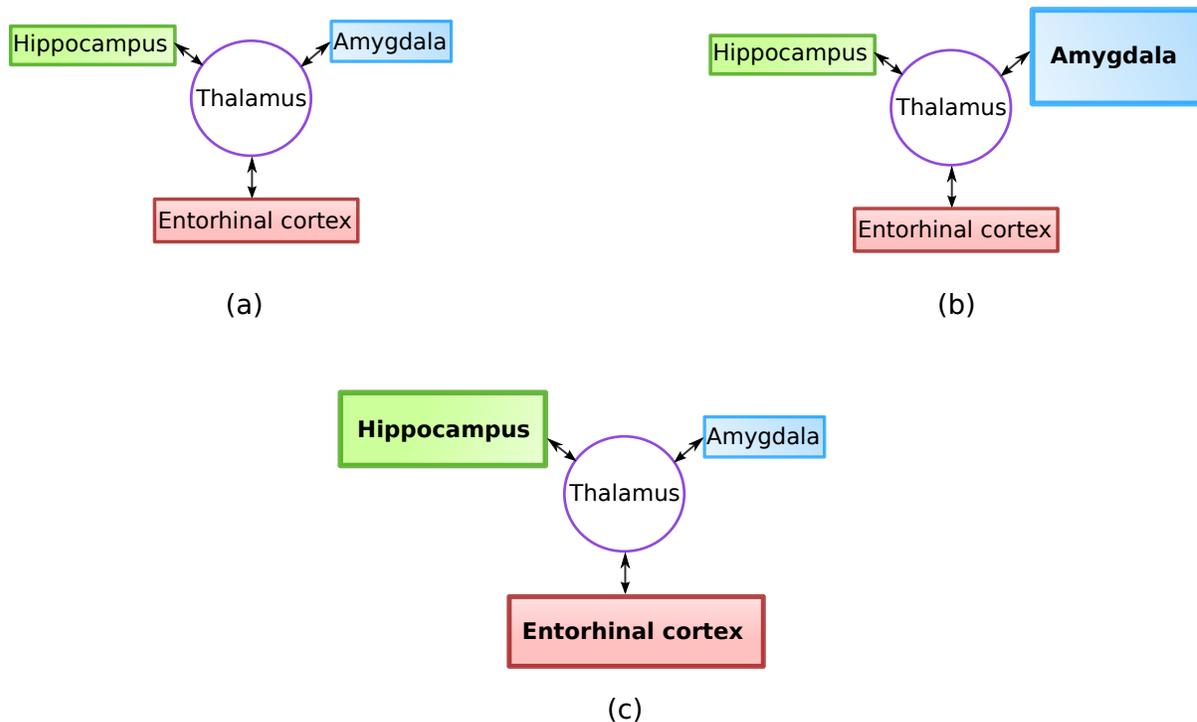


Figure 1.3: **Hypothesis of multiple independent generators in TLE.** a) a situation where all three structures contribute equally to the seizure. b) Theoretical amygdala-dominant seizure. c) A seizure in which the amygdala would play a little role in relation to the other structures.³

There are several pathologies in the temporal lobe that can be responsible for seizures. These include tumours, vascular and developmental malformations, cortical dysplasia, tubers,

³Adapted from *Epilepsy & Behavior* 14 (2009), Bertram, Edward H., *Temporal lobe epilepsy: Where do the seizures really begin?*, 32-37, © 2015, with permission from Elsevier.

trauma, infections, genetic syndromes, and metabolic disorders. It is presumed that the seizure onset will vary according to the location of the pathology [11]. On the other hand, in the absence of an established aetiology, it has been theorized that seizures originate as the result of the interaction of multiple regions and that in reality there is no single seizure onset focus (Figure 1.3).

There has also been some speculation about the function of *subcortical structures* in ictal discharges (during seizures). For example, experiments in animal models have been conducted showing that the thalamus can play a role in the regulation of temporal lobe seizures [11]. The origin of seizures has also been studied using post-surgical resections and in *post-mortem* cases pointing to hippocampal sclerosis (cellular loss, atrophy and proliferation of glial cells in the hippocampus) as the most common TLE pathology. Other regions where abnormal changes have been found are the amygdala and entorhinal cortex.

The detection of seizures *in vivo* is a clinical challenge that is currently being addressed by two important technologies: electroencephalography and neuroimaging.

1.1.2 Role of Electroencephalography

Encephalography is a technique that measures the electrical activity in the cortex using electrodes placed on the scalp. EEG cannot record electrical activity from individual neurons for two reasons: neurons are very small, and their signal is very faint. However, electrical recordings are still possible due to the fact that pyramidal neurons (one of the largest type $\sim 20\mu\text{m}$) are physically aligned, have the same polarity and many fire in synchrony (Figure 1.4).

Electrodes are organized in an array covering the scalp, with the most common array being the international 10-20 system (Figure 1.5). This array can be configured to measure voltage differentials between electrode pairs (bipolar) or between each electrode and a common ground (common reference). The interpretation of the results vary according to this configuration.

EEG plays an important role in the diagnosis of epilepsy. First, the presence of seizures can be confirmed. It is estimated that between 4% and 10% of patients have psychogenic ⁴

⁴i.e. a *psychological* rather than a *physiological* cause

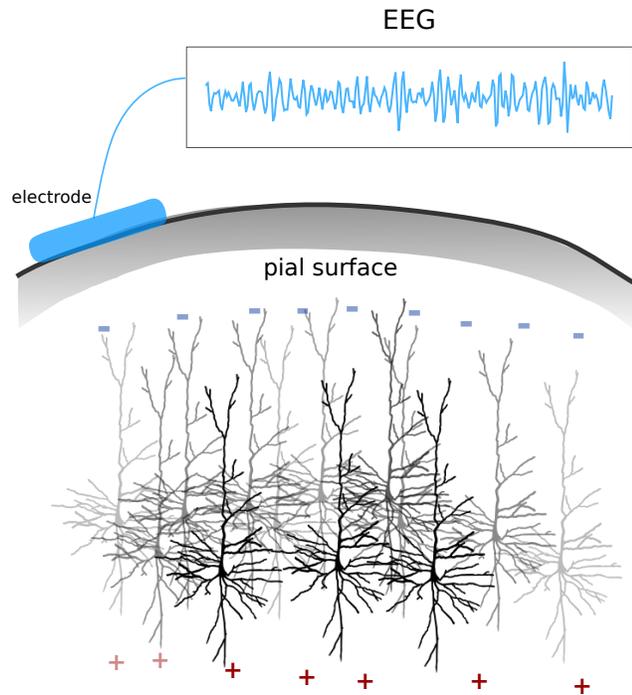


Figure 1.4: **EEG principle.** The electrical field generated by similarly oriented pyramidal neurons can be detected by a scalp electrode. Pyramidal neurons possess a pyramidal-shaped soma and two distinct dendritic trees: apical and basal.

disorders instead of epilepsy [12], thus EEG can discriminate a psychogenic seizure from an epileptic seizure by establishing a correspondence between the clinical semiology ⁵ and the recorded EEG signals. Second, EEG can determine if seizures are generalized or focal, and it can also help investigate whether the patient has a specific epilepsy related syndrome. These findings can help determine a clinical course of action including specifying the type of AEDs and determining whether the patient is a candidate for surgery.

Nevertheless, EEG has several limitations. During ictal periods, EEG readings are marred, and signals from the scalp muscle, motion and electrode artifacts can limit interpretation [4]. Moreover, EEG has a low sensitivity for the detection of inter-ictal ⁶ epileptogenic discharges (IED) in temporal lobe areas (between 29% and 58%) [12–14]. To increase the yield, more

⁵sign and symptoms of seizures

⁶**inter-ictal**: interval between seizures

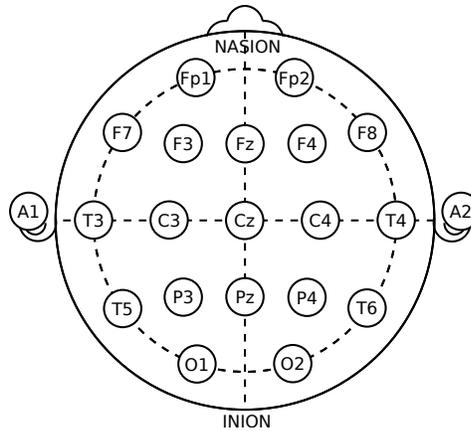


Figure 1.5: **10-20 electrode array**. T: temporal, F:frontal, Fp: frontal polar, C:central reference, O:occipital, P:parietal, z:midline; even numbers: right hemisphere, odd numbers: left hemisphere; nasion: between the eyes above the bridge of the nose, inion: lowest point of the skull on the back of the head. Nasion and inion are landmarks for positioning of the array.

electrodes over the temporal region, also known as Silverman electrodes [15], can be added to the standard 10-20 system. Another approach is long term monitoring, where waking and sleeping inter-ictal readings are combined, increasing sensitivity to 80% [12].

Intracranial EEG

The implantation of subdural or deep electrodes is indicated when a patient is being considered for surgery, and the readings of scalp EEG have been inconclusive. While scalp electrodes capture small signals in the microvolt range, subdural electrodes record signals with a larger amplitude between 10 mV to 20 mV, thus being more sensitive. Additionally, subdural electrodes are less sensitive to muscle and motion artifacts, have a good temporal resolution and can be placed in zones near the mesial temporal region to obtain more precise measurements (Figure 1.6). Despite being invasive, the recordings obtained with subdural electrodes are more detailed and better localize seizures. Also, subdural readings can provide insight into cases where scalp electrodes report bilateral ictal or inter-ictal abnormalities, or where the seizure onset cannot be determined [14].

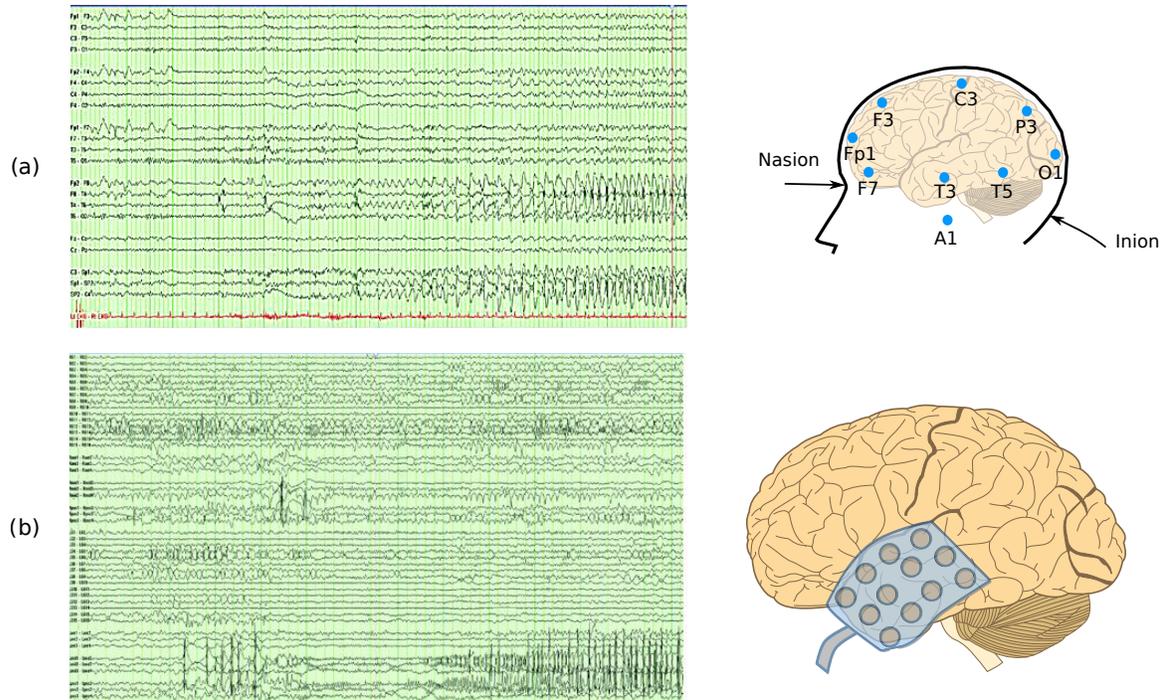


Figure 1.6: **Temporal lobe seizure captured with EEG.** a) seizure recorded using the traditional 10-20 system b) seizure recorded using subdural electrodes (24 channels per lobe). Subdural channels have better localization power but surgical implantation is required.⁷

1.1.3 Role of Imaging

Structural MRI

MRI measures how protons contained in tissue interact with external electromagnetic fields. Such interaction is determined by the proton density (ρ), and the T1 and T2 magnetic properties of tissue. Defined according to the orientation of the main magnetic field B_0 in the MR scanner, T1 measures magnetic recovery along B_0 (regrowth of longitudinal magnetization), while T2 measures magnetic relaxation orthogonal to B_0 (loss of transverse magnetization).

The type of imaging sequence and the particular set of acquisition parameters determine the relative contribution of T1 and T2 to the image, generating T1-weighted, T2-weighted, or proton density weighted images.

In the study of TLE, structural MR images allow for the identification of [16]:

⁷Adapted under the Creative Commons Attribution License from Epilepsy Research and Treatment 2012, “Electroencephalography in Mesial Temporal Lobe Epilepsy: A Review”, Javidan M., © 2012.

- Tumours: meningiomas, gangliomas (more common in children), low-grade astrocytomas, dysembryoplastic neuroepithelial tumours (DNET), and cavernomas.
- Hippocampal sclerosis: also known as mesial temporal sclerosis (mesial temporal sclerosis (MTS)). The most common finding in patients undergoing surgery and also the most common cause of partial complex seizures. MTS is characterized by apoptosis of pyramidal neurons in the hippocampus (cornu ammonis and dentate gyrus), hippocampal reorganization and changes in energy metabolism [17].
- Malformations of cortical development: focal cortical dysplasia (FCD) is one of the most common cortical developmental malformations. These correspond to groups of neurons that failed to migrate to the proper formation *in utero*. Another cortical development malformation is tuberous sclerosis, a genetic condition causing benign tumoural growth.

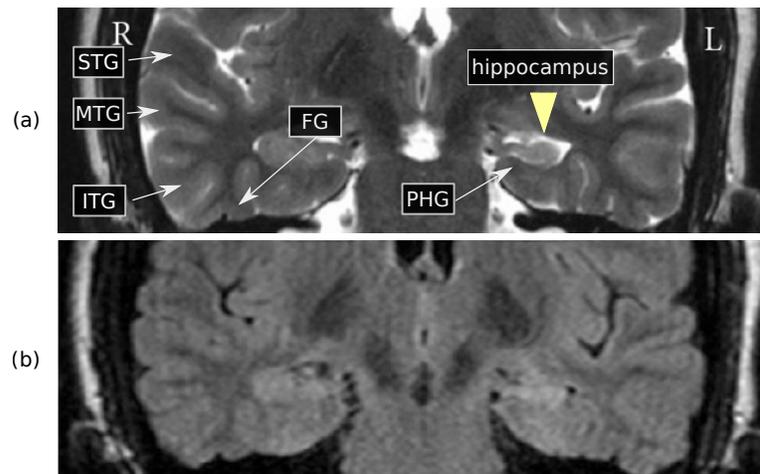


Figure 1.7: **Structural imaging in TLE** This figure shows a preoperative MR study for a patient with left mesial temporal sclerosis (indicated by yellow arrow). STG: superior temporal gyrus, MTG: middle temporal gyrus, ITG: inferior temporal gyrus, FG: fusiform gyrus, PHG: parahippocampal gyrus. a) Coronal T2-weighted image; b) FLAIR image.⁸

Fluid attenuation inversion recovery (FLAIR) is a sequence used to diagnose TLE. In FLAIR images, cerebrospinal fluid (CSF) is nulled by selecting an *inversion time*, TI, such

⁸Adapted under the Creative Commons Attribution License from Epilepsy Research and Treatment 2012, “Surgical Techniques for the Treatment of Temporal Lobe Epilepsy”, Al-Otobi et al., © 2012.

that $TI = \ln(T1_{CSF})$. This improves the detection of lesions in the boundary between gray matter and CSF(Figure 1.7).

The identification of lesions with structural MRI can determine the clinical course of action: FCD and MTS patients frequently become drug-resistant, and these lesions are usually removable with surgery. In contrast, neoplasms (tumours) usually respond well to AEDs [17] and surgery is not always required. Likewise, the localization, extension, and type of lesion determine whether or not it can be surgically removed.

In cases where MRI lesions are absent (also known as MR negative, or non-lesional) neurologists turn to subdural EEG, to confirm epileptogenicity. Placing subdural and deep electrodes is invasive, has potential risks, and needs to be done with precision. Hence, MRI is used to guide positioning, and to validate the anatomical distribution of electrodes.

MRI can also be a valuable post-surgical resource, to determine the reasons for operative failure or complications, and to monitor the recurrence of lesions such as tumours.

Functional MRI

Functional MRI tracks local changes in blood flow in response to external stimuli (i.e. visual, auditory, verbal), and is used pre-surgically to identify eloquent regions in the cortex.⁹ This is particularly important as language and memory areas commonly overlap with TLE lesions, and during surgery these regions must be avoided to minimize cognitive deficits.

Traditionally, the localization of eloquent areas has been performed with the Wada test, which consists of injecting sodium amobarbital into the carotid arteries, one hemisphere at a time. This sedative has the effect of diminishing language and memory functions in the hemisphere where is injected. During the test, the patient is engaged in a series of tasks to assess memory and language functions, thus assessing these functions in the hemisphere that is not anesthetized. This test is invasive with possible complications including stroke and death [18].

Though there has been a good agreement between Wada and fMRI studies, the applicability of fMRI to language and memory mapping is an area of active research. For instance, it

⁹The *eloquent cortex* refers to regions in the cortex that, if removed, will result in loss of sensory perception, speech and motor functions

has been speculated that neurological disorders, such as TLE, can modify networks that are commonly activated in the healthy population therefore making current stimuli paradigms less effective in the presence of such disorders [18, 19].

PET, SPECT and MRS

In addition to structural and functional findings, neuroimaging provides valuable diagnostic information regarding abnormal brain metabolism and chemical composition. In this section three imaging techniques are presented. Similar to structural and functional MRI, these techniques have been used in the characterization and lateralization of lesions, being used independently or jointly by clinicians to validate EEG and MRI findings. A good level of agreement among techniques is usually correlated to better surgical outcomes. These techniques are: PET, SPECT and MRS.

Positron emission tomography (PET) is a technique that generates images of metabolic processes in the body, by employing positron-emitting radioisotopes attached to biologically active molecules, also known as *radiotracers*.

A common radiotracer used in studying TLE is fluorodeoxyglucose (FDG)¹⁰ which allows the visualization of glucose metabolism. Reduced FDG uptake (glucose hypometabolism) has been observed during inter-ictal periods in patients with TLE (Figure 1.8) [20, 21]. PET-FDG is generally used in MR negative patients, or to analyze the spread of seizure patterns providing additional information for the placement of subdural electrodes. Moreover, PET has a high specificity for MTS, as MTS is associated with hypometabolism in the hippocampus, amygdala, entorhinal cortex and temporal pole [22]. Though not routinely employed, ictal PET provides evidence of hypermetabolism in the epileptogenic zone with suppression of metabolically activity in the surrounding areas [22, 23].

Single-photon emission computed tomography (SPECT) is similar to PET in that radiotracers are used to localize metabolic activity. Unlike PET, gamma photons are emitted directly by the radiotracer; therefore SPECT detects “single”-photons. SPECT radiotracers are

¹⁰Formally known as ¹⁸F-FDG, a glucose analog with the positron-emitting radioactive isotope fluorine-18

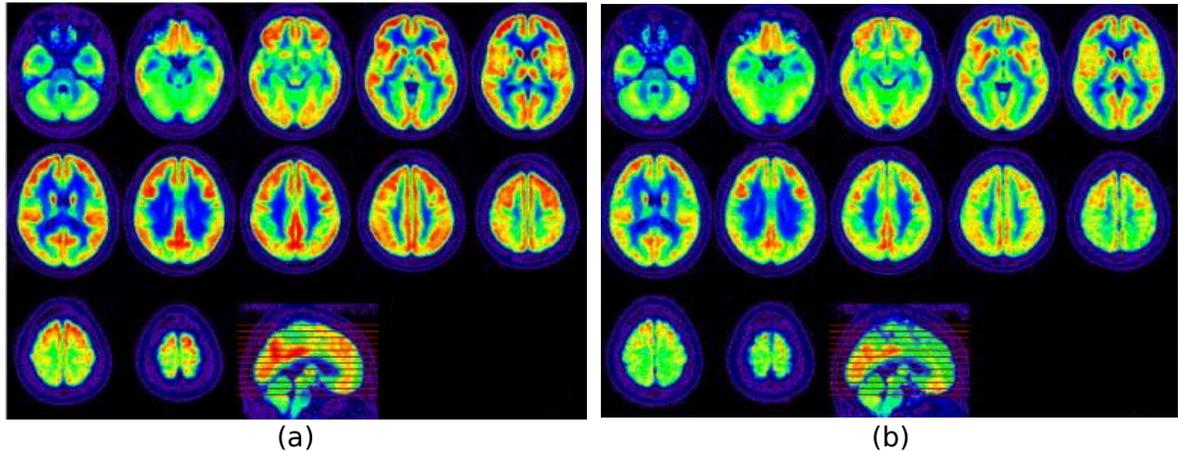


Figure 1.8: **PET-FDG**. Visual assessment of the mean images of normalized FDG-PET across (a) healthy volunteers and (b) TLE patients, in Nelissen et al. study showing a striking hypometabolism in the fronto-parietal lobes in patients compared to control subjects.¹¹

longer-lived and easier to manufacture which makes SPECT a more economical alternative than PET. However, SPECT images have a much coarser resolution than PET. Technetium-99m-hexamethylpropylene-amine-oxime (99mTc-HMPAO) [20, 24] is a SPECT radiotracer used for measuring cerebral blood flow (perfusion). This radiotracer visualizes ictal hyperperfusion that remains some time post-ictally [23]. Thus, its injection right after a seizure can help identify epileptogenic zones. The sensitivity of this method is increased by comparing ictal against inter-ictal perfusion images [22].

Magnetic resonance spectroscopy (MRS) enables chemical composition analyses through the study of MR spectra. In TLE, MRS is used to measure levels of N-acetylaspartate (NAA) which is a marker of metabolic active neurons. Reduced NAA indicates neuronal loss and/or metabolic dysfunction [17]. In contrast, creatinine (Cr) and choline (Cho) are present at higher levels within glial cells in TLE patients. Nevertheless, MRS is not commonly used due to its technical challenges and lack of widespread availability [22].

¹¹Adapted from NeuroImage, 32 (2006), N. Nelissen et al., *Correlation of inter-ictal FDG-PET metabolism and ictal SPECT perfusion changes in human temporal lobe epilepsy with hippocampal sclerosis*, 648-695, © 2015, with permission from Elsevier.

1.1.4 Surgery: clinical challenges and current techniques

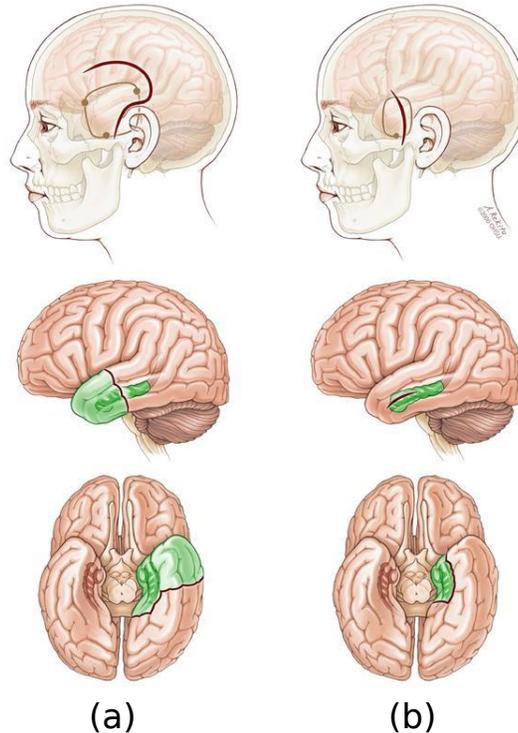


Figure 1.9: **Temporal lobe surgery** a) anterior temporal lobectomy (ATL); b) selective amygdalohippocampectomy (SAH). The first row shows the craniotomy, the second and third rows show the target tissue.¹²

EEG and imaging information are necessary for planning the surgical approach in those patients that require it: between 15% and 30% of patients with focal epilepsy become drug resistant [17] suffering uncontrolled seizures that diminish their quality of life. For them, surgery is an option only if the epileptogenic focus has been clearly identified. Otherwise, invasive subdural and/or deep electrodes are necessary to localize seizures. Nevertheless, when intracranial EEG reveals multiple epileptic foci or provides insufficient information, temporal lobe surgery cannot be performed, and in some cases sectioning the corpus callosum (corpus callosotomy) is the only palliative measure available [17].

When surgery is viable, anterior temporal lobectomy (ATL) or selective amygdalohippocampectomy (SAH) are performed, with the former being the most common. In ATL, the

¹²Reproduced under the Creative Commons Attribution License from Epilepsy Research and Treatment 2012, “Selective Amygdalohippocampectomy”, Spencer D. and Burchiel K., © 2012.

anterior portion of the temporal lobe (including mesial structures) is resected *en-bloc*, whereas in SAH, only the mesial structures (amygdala, hippocampus) are removed (Figure 1.9). SAH can only be used when the seizure onset zone has been identified in the mesial temporal region only. Some risks associated with temporal lobe surgeries are motor, visual, language and memory deficits, and an increased risk of depression and suicide [22]. However, success rate is good, with low morbidity (about 11%) [22].

1.2 MRI research in Temporal Lobe Epilepsy

1.2.1 Quantitative Relaxometry

Quantitative relaxometry (Q-MRI) refers to the measurement of biophysical properties through decoupling of the different contrast mechanisms that contribute to the overall MR signal [25].

T1- and T2-weighted images

Prior to delving into the acquisition of quantitative T1 and T2 maps, we will take a look at how voxel intensities are determined in MRI. In spin echo sequences, the voxel intensity, S , can be characterized as:

$$S \propto \rho(1 - \exp[-TR/T1]) \exp[-TE/T2] \quad (1.1)$$

with the acquisition parameters, TR and TE, being the repetition and echo times respectively. Short TR and short TE generate images where the intensity contribution from T1 is larger (T1-weighted), whereas long TR and long TE produce images governed by the T2 parameter (T2-weighted). On the other hand, the estimation of proton density weighted images (ρ) is obtained by setting a long TR and a short TE. In Equation 1.1 a long TR makes the term $1 - \exp[-TR/T1]$ close to 1, and with a short TE, the term $\exp[-TE/T2]$ close to 1 as well, therefore making the signal proportional to proton density.

Acquisition parameter values are selected on the basis of improving radiological assessment, producing tissue contrasts that enable the physician to compare tissues and locate abnor-

malities. Thus, voxel intensities have meaning only with reference to other voxels and to the amount of contrast they provide for visual inspection.

Obtaining quantitative T1 and T2 maps

T1 and T2 can be estimated by approximating the MR signal with exponential equations for very long TRs (repetition times). For T1, after applying a 180° *inversion pulse* and a long TR, the *longitudinal magnetization recovery* M_z can be modeled as:

$$M_z(t) = M_0 \left(1 - 2 \exp \left[-\frac{t}{T1} \right] \right) \quad (1.2)$$

which can be linearized as:

$$\ln \left[\frac{1}{2} \left(1 - \frac{M_z(t)}{M_0} \right) \right] = -\frac{t}{T1} \quad (1.3)$$

hence, T1 can be estimated as the the inverse of the logarithmic slope in Equation 1.3. As for T2, after a 90° RF pulse, and a long TR, the *transverse magnetic decay* M_{xy} can be approximated as:

$$M_{xy}(t) = M_0 \exp \left[-\frac{t}{T2} \right] \quad (1.4)$$

and after algebraic manipulation presented as a liner form:

$$\ln \left[\frac{M_{xy}(t)}{M_0} \right] = -\frac{t}{T2} \quad (1.5)$$

To obtain a line, a minimum of two measurements are needed. In terms of the MR sequence, this means scanning for at least two echo times (TE). Though simple, these methods are not widely used in clinical settings due to the lengthy scanning times: T1 mapping requires an inversion recovery pulse, whereas T2 mapping needs multiple echos, both in the context of

spin echo sequences¹³. Nevertheless, the opportunity of studying pathologies in terms of T1 and T2 properties has spurred interest to the development of time-efficient sequences.

Accelerated methods for T1 and T2 measurement

Possibly one of the most common methods to accelerate the acquisition of T1 maps is the Lock and Locker method [26]. This method also requires an inversion pulse (180°). After this pulse, the recovery of longitudinal magnetization is measured with a train of very small-angle RF pulses that disturb minimally the magnetization recovery along B_0 , making possible the estimation of the magnetization recovery curve with only one inversion pulse [25, 27]. Another method is the Driven Equilibrium Single Pulse Observation of T1 (DESPOT1) [28]. This method uses a spoiled gradient recalled echo (SPGR) sequence to acquire images at several small flip angles (α). The SPGR signal (S_{SPGR}) can be written as:

$$\frac{S_{SPGR}}{\sin \alpha} = \frac{S_{SPGR}}{\tan \alpha} \exp\left[-\frac{TR}{T1}\right] + \rho(1 - \exp\left[-\frac{TR}{T1}\right]) \quad (1.6)$$

from which T1 can be calculated as a function of α . Thus, this method requires at least two distinct flip angles (two SPGR images). An immediate consequence of this approach is that the estimation will be only as good as the certainty of the assumed flip angle, meaning that the uniformity of the RF field (B_1) is a relevant factor. In this work we use the High-speed Incorporation of RF Field Inhomogeneities DESPOT1 or DESPOT1-HIFI, which adds an extra sequence, a inversion recovery SPGR (IR-SPGR) used to improve the estimation of T1 and to compensate for B_1 inhomogeneities (Figure 1.10a). DESPOT1-HIFI allows a whole brain to be scanned with a voxel size of 1 mm^3 in around 10 minutes [29].

Similar to the idea of a variable flip angle for the estimation of T1, a variable flip angle fully balanced Steady State Free Precession (bSSFP) [30] can be used for the fast acquisition of T2 maps. Like SPGR, SSFP is also a gradient echo sequence with small flip angles. Unlike

¹³In spin echo sequences measurable signals or, *echos*, are generated by refocusing the transversal magnetization with an electromagnetic pulse

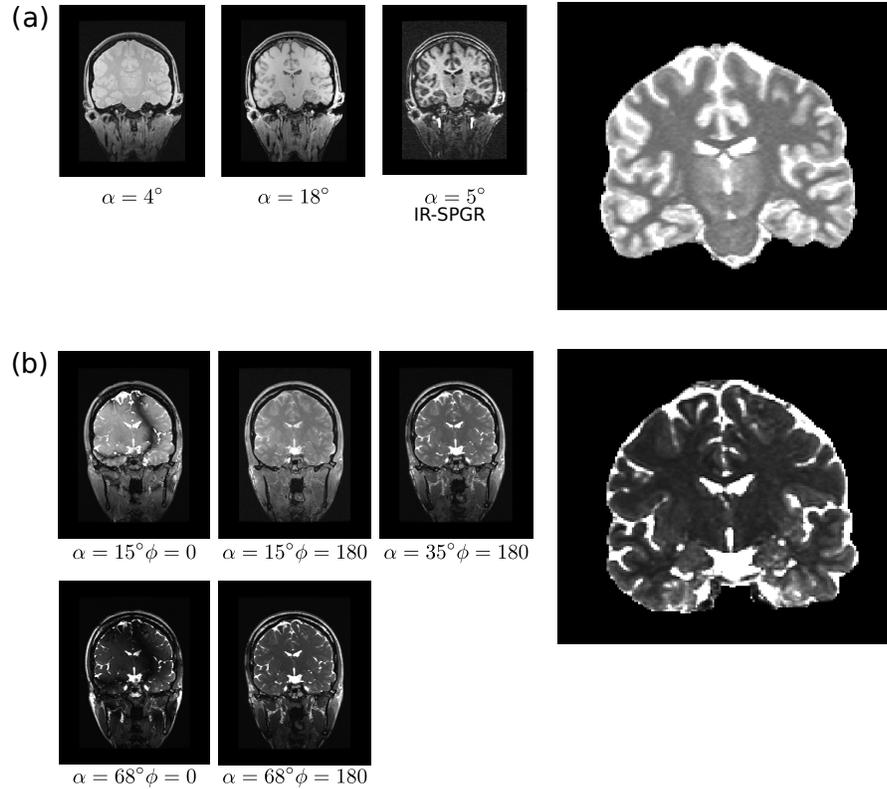


Figure 1.10: **Quantitative Relaxometry.** a) Estimation of a T1 map using DESPOT1-HIFI; b) Estimation of a T2 map using DESPOT2-FM.

SPGR however, the repetition times in SSFP are very short, and there is no spoiling¹⁴ because we are actually interested in the transversal magnetization. With a very short TR ($<10\text{ms}$), the RF pulses are close together and the MR signal will never completely decay, thus maintaining phase coherence in the transverse plane. Collecting SSFP images over a range of flip angles produces a signal curve that is function of both T1 and T2 [31]. Having obtained the T1 map with any of the aforementioned methods, the estimation of the T2 map follows. This is the basis for the DESPOT2 technique which is used in this work to estimate T2 maps (Figure 1.10b). Additional methods to generate T2 maps using fast spin echo (FSE) and magnetization-prepared sequences have also been reported in the literature [25].

Q-MRI has proven to be a sensitive imaging technique for the detection of TLE related pathologies. T1 and T2 values appear increased in the temporal lobe, particularly in the hemisphere containing the seizure focus. Additionally, a correlation between T2 hippocampal hy-

¹⁴**spoiling:** destruction of the transverse magnetization using gradients

perintensities and the presence of mesial temporal sclerosis (MTS) has been established [32–35].

1.2.2 Diffusion tensor imaging

Diffusion tensor imaging (DTI) enables the examination of morphological integrity in the brain by analyzing how water molecules diffuse in tissue. This is particularly relevant for imaging white matter where diffusion is restricted due to the myelin sheath around neuronal axons.

Imaging principle: diffusion weighted imaging

Water diffusion in tissue is measured using a modified spin echo sequence that is sensitive to the Brownian motion of water, followed by a fast readout (echo-planar imaging). This kind of imaging sequences are known as *diffusion weighted imaging* (DWI).

In a spin echo sequence, spins are tipped onto the transverse plane using a RF pulse. After the RF pulse is removed, spins start de-phasing and a *refocusing RF pulse* and gradient are applied to reverse this process and to obtain a signal during the readout time (echo time, TE). A pair of identical gradients, known as *diffusion gradients*, can be added to the sequence: one immediately before the refocusing pulse and the other right after it (Figure 1.11a). The pre-refocusing gradient changes the phase of spins in each voxel, and if there is no diffusion, the post-refocusing gradient reverses this phase, canceling the effect of the first gradient. Thus, during the readout, no net effect is recorded (Figure 1.11b). However, if there are free water molecules in a voxel, these can leave the voxel at any time generating a measurable net effect in phase when the post-refocusing gradient is applied. This effect is proportional to the free motion of these water molecules, or in other words to their diffusion (Figure 1.11c). The b-factor controls the amount of contrast in DWI. The higher the b-value, the higher the contrast due to diffusion. The b-factor depends of the shape and timing of the diffusion gradients employed as shown in Figure 1.11a.

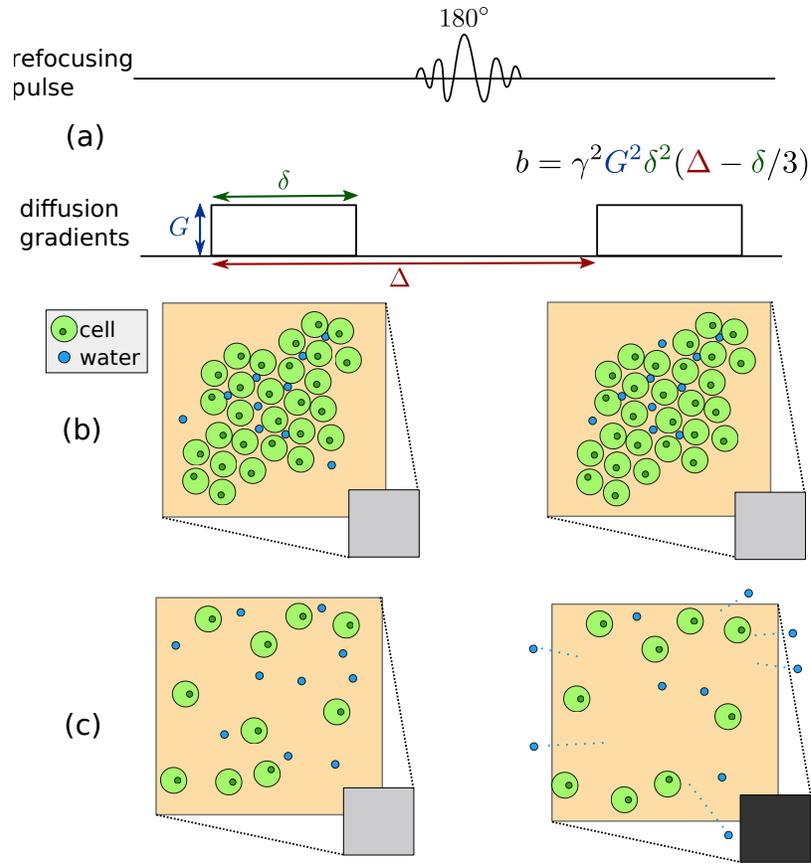


Figure 1.11: **Measuring diffusion.** a) diffusion gradients. The b -factor characterizes a diffusion sequence and relates to the shape and spacing of the diffusion gradients; b) low diffusion: Brownian motion of water molecules is restricted and the net effect of the diffusion weighting is minimal; c) high diffusion: there is a drop in signal due to the free motion of water molecules outside of the voxel, which changes the net effect of the diffusion gradients.

The diffusion tensor

A robust three dimensional characterization of the diffusion process in every voxel can be achieved by combining the information provided by the individual diffusion weighted images. Each DWI describes the magnitude of the diffusion process in a particular *spatial orientation* given by the gradients employed. This information can be encoded in the *diffusion tensor*:

$$D = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} \quad (1.7)$$

Due to the symmetry of the diffusion tensor ($D_{ij} = D_{ji}$) a minimum of 6 orientations (6 DWI) must be acquired. As more orientations are recorded, the estimation accuracy will increase accordingly.

Technical considerations

Obtaining optimal diffusion weighted images greatly depends on the ability to generate strong, rapid gradient pulses. However, when diffusion gradients are switched on and off, the resulting time-varying magnetic field induces *eddy currents* in the scanner hardware [36] resulting in shift and shear distortions in the images. Thus, eddy-current correction must be addressed in these images prior to the estimation of the respective diffusion tensors. Eddy-current correction may be implemented with a first-order affine transform to an undistorted, non-diffusion weighted, reference volume of the same subject. Similarly, EPI-related artifacts can be minimized with parallel acquisition techniques, and non-linear registration alternatives [36–38].

Mean Diffusivity and Fractional Anisotropy

An eigenanalysis of the diffusion tensor provides information on the three-dimensional diffusion process occurring on each voxel level. The rate at which diffusion occurs is represented by the mean diffusivity (MD), and it is defined as:

$$MD = (\lambda_1 + \lambda_2 + \lambda_3)/3 \quad (1.8)$$

where $\lambda_1, \lambda_2, \lambda_3$ correspond to the three eigenvalues of the diffusion tensor. High MD values reflect regions where diffusion is unrestricted such as the ventricles. In contrast, the MD in white matter is lower (Figure 1.12).

Another measurement of interest is the fractional anisotropy (FA) which indicates spatial restriction in the diffusion process, ranging from 0 (isotropic diffusion) to 1 (ideal linear diffusion), and it is given by:

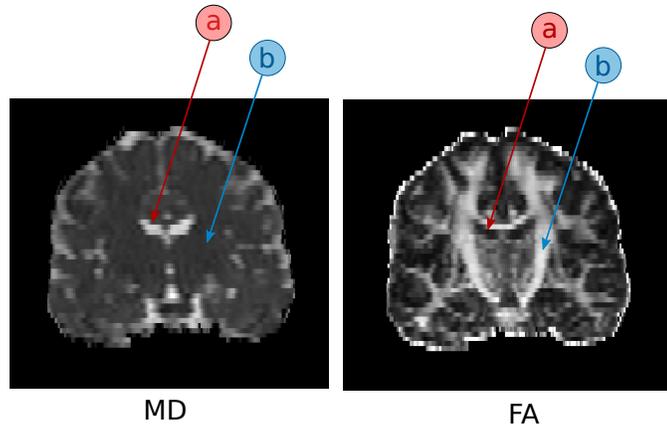


Figure 1.12: **Mean diffusivity (MD) and fractional anisotropy (FA)**. a) water diffusion in the ventricles is unrestricted, thus MD is high and FA is low. b) water diffusion in the corticospinal tract is restricted by the myelin sheath covering the axon bundles, so the diffusion is highly anisotropic (high FA), while the MD image indicates that the overall diffusion is low.

$$FA = \sqrt{\frac{3}{2}} \sqrt{\frac{(\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \quad (1.9)$$

White matter is characterized by its high FA, whereas diffusion is more isotropic in regions containing CSF (Figure 1.12).

In the study of TLE, it has been shown that there is a pattern of ipsilateral FA loss in the temporal white matter (uncinate fasciculus, arcuate fasciculus, inferior longitudinal fasciculus) accompanied by MD increase in the hippocampus and temporal gray matter [39, 40] though minor contralateral FA decline has also been reported [41].

MD and FA in the study of TLE

Similarly to Q-MRI studies, comparisons of DTI abnormalities between MTS and non-MTS patients, have shown differences in morphological change patterns with MTS patients having an increased ipsilateral MD signal in the hippocampus, parahippocampal gyrus and frontoparietal regions, while no MD alterations were found in non-lesional subjects [39, 42]. Similar results have been obtained in FA analyses, where white matter abnormalities are more predominant in the MTS group compared to the non-lesional group [43, 44].

1.3 Imaging Analysis Techniques

1.3.1 Region-of-interest techniques (ROI)

The analysis of regions of interest (ROI) in MRI images is one of the first computational approaches available for the study of TLE. These techniques rely on the expertise of a user to perform manual or interactive segmentation [45–47], or in algorithms that employ anatomical atlases to guide the segmentation [48–52].

Given that patients with MTS generally present loss of hippocampal gray matter, and often loss of internal architecture ipsilateral to the lesional side, ROI techniques have focused on assessing hippocampal volume [53], shape and pose [46, 54]. ROI methods have identified volume reductions that are not detectable by visual inspection, thus increasing the detection of hippocampal sclerosis in about 28% of patients [52]. Similarly, salient ipsilateral volume reductions in the putamen, thalamus, amygdala, entorhinal and perirhinal cortices have also been reported in TLE patients [45, 50, 55].

ROI techniques also examine regional mean intensities. T2 signal loss has been described in the hippocampus, amygdala and thalamus and other temporal structures, mainly ipsilateral to the seizure onset [34]. Likewise, ROI analyses of DTI-derived images, show bilateral reduction of the FA signal in the hippocampus, parahippocampal, temporal white matter, frontal lobe white matter, internal capsule, corpus callosum and brain stem ROIs, with predominance of changes in mesial temporal lobe patients, particularly in the left MTS group [47, 50].

1.3.2 Voxel-based statistical techniques

The premise of voxel-based techniques is to establish a voxel-to-voxel correspondence among subjects through non-linear image registration to a standard brain template. Once all images are co-registered, differences can be obtained by performing voxel-wise statistical tests. The majority of these techniques employ group-wise designs, with statistics revealing differences among two subject groups. In the case of TLE, these groups are usually patients and controls; patients with left or right temporal lobe seizure onset (L-TLE, R-TLE); and patients with MTS

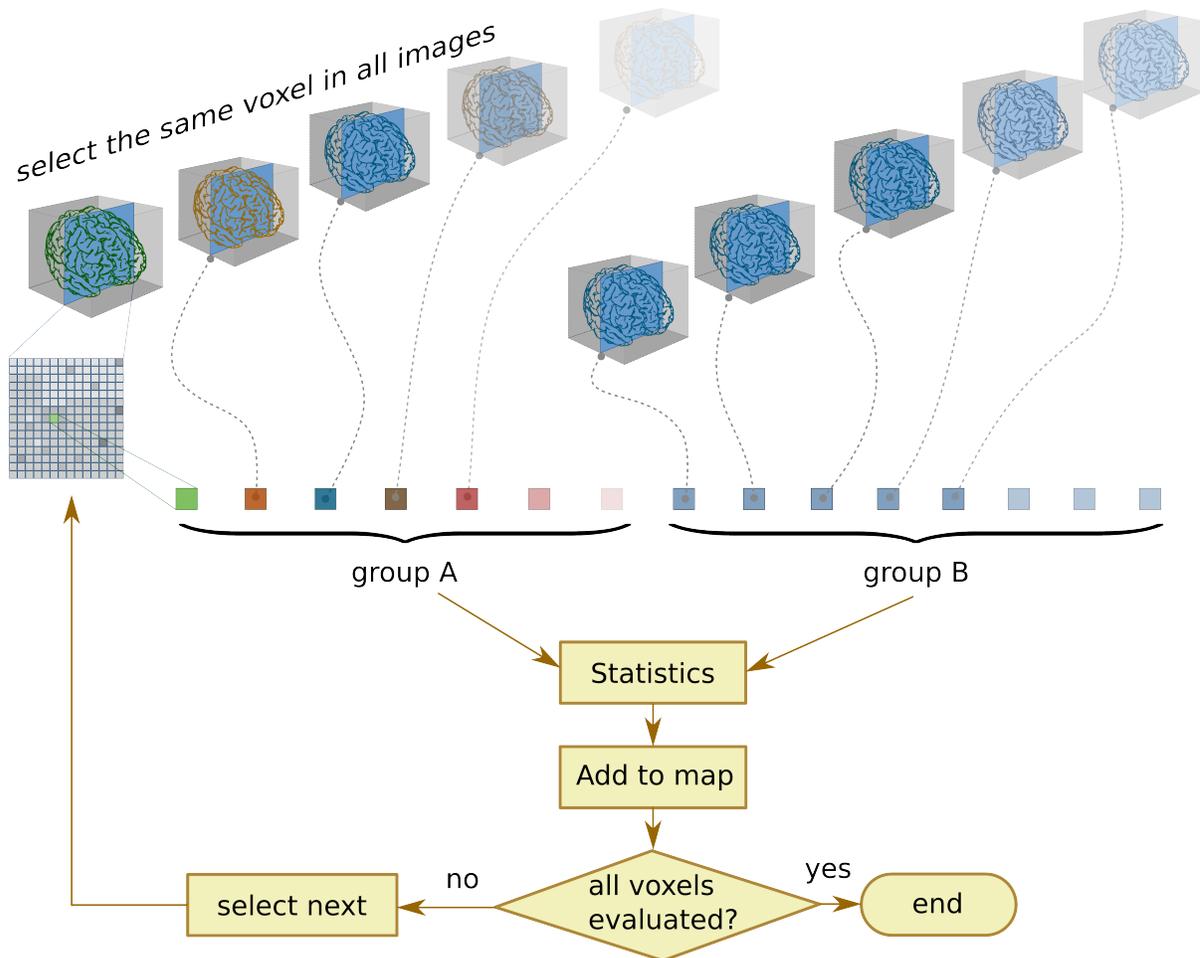


Figure 1.13: **Voxel-based statistical techniques.** These techniques rely on a good spatial normalization to compare groups of images at the voxel scale. A statistical map is obtained by performing voxel-wise statistics. A correction for multiple comparisons is usually implemented.

vs. patients with no MTS. Individual analyses are attainable by comparing one subject against a group of healthy controls but these designs are limited as we will discuss in this section.

Voxel-based morphometry (VBM)

Voxel-based morphometry (VBM) [56] is a technique that detects local tissue differences in a population by performing voxel-wise comparisons using T1-weighted images. In particular, VBM detects changes in gray matter (GM), white matter (WM) and CSF. The standard VBM method involves four steps: (i) spatial normalization to an anatomical template; (ii) segmentation of GM, WM, and CSF tissues; (iii) spatial smoothing that compensates variability in

individual gyral anatomy; and (iv) statistical analysis [57].

An optimized VBM method has been proposed by Good et al. [58]. In this variation of the technique, optimized spatial normalization parameters are acquired as follows: (i) the original image is affine-registered to the anatomical template, and GM and WM segmentations are obtained; (ii) The GM and WM segmentations are registered to the template generating optimal normalization parameters which are then applied to the original image. The rationale behind this optimization is that in the original technique, scalp fat, dural venous sinuses and other structures can be misidentified as gray matter voxels, hence, a registration driven by pre-segmented GM and WM regions can be more accurate.

An additional step known as *modulation* is often included in optimized VBM. This step multiplies GM voxels by the Jacobian determinants which are estimated during spatial normalization. This is to adjust GM volumes according to the spatial deformation undergone to match the anatomical template. To illustrate, if one subject's hippocampus has half of the volume of that of the template, then its volume will be doubled during spatial normalization, consequently doubling the number of voxels labelled as gray matter. Similarly, if another cortical structure is twice the size of the respective structure in the anatomical template, then the number of voxels will be halved after registration. Multiplying the registered volume by the Jacobian determinant has the effect of preserving the total amount of gray matter, thus enabling the statistical comparison of gray matter volumes.

Optimized VBM has been compared with standard VBM in the study of temporal lobe epilepsy. It has been shown that optimized VBM delivers complimentary volume reductions (modulated results) that match tissue concentration reductions (unmodulated results) identified with the standard technique [59] and that it may reveal subtle changes in gray matter volume that are not otherwise captured [60].

The most common TLE-related finding in VBM studies is the loss of gray matter in the hippocampus ipsilateral to seizure onset, and in some cases contralaterally [53, 57] in MTS patients. Other structures such as the thalamus, amygdala, parahippocampal gyrus (entorhinal and perirhinal cortices), fusiform gyrus, inferior temporal, middle temporal, and superior

temporal gyrus, temporal pole and fornix have shown asymmetric changes with a predominant GM loss ipsilaterally [57]. Gray matter loss has been more commonly found in MTS patients than non-MTS patients [59]. Analogously, evidence of WM loss adjacent to mesial temporal lobe (MTL) structures ipsilateral to seizure onset (and in some cases contralaterally) has been reported in the literature [53].

VBM has also been used to detect focal cortical dysplasia on individual patients, in FLAIR images [61], and in T1- and T2-weighted images [62]. In these studies, patient images are compared individually against a group of healthy individuals, and the resulting statistical maps are able to pinpoint the location of the FCD in most cases. Nevertheless, there is evidence for not using VBM for individual patient assessments. VBM is not sufficiently sensitive to detect focal pathology in individuals with relatively subtle brain changes [57], such as cortical neuronal loss and hippocampal sclerosis [63] and, beyond the realm of TLE, it has been shown that VBM is not an adequate stand-alone technique for detecting focal lesions [64].

Voxel-based relaxometry (VBR)

The idea of performing voxel-wise comparisons in a population has been extended to quantitative relaxometry. This technique, known also as voxel-based relaxometry (VBR) [33], follows a similar processing pipeline to the standard VBM technique, with the exception that there is no segmentation. Instead, images are: (i) spatially normalized to a template; (ii) smoothed using a Gaussian kernel; and (iii) statistically analyzed. In VBR, spatial normalization occurs by proxy of a T1-weighted or T2-weighted image: The transformation parameters obtained from registering such an image to the anatomical template are applied to the quantitative maps. VBR results are reported in terms of *relaxation times* (T_1 , T_2), but it is also common to find studies referring to *relaxation rates* ($R_1 = 1/T_1$, $R_2 = 1/T_2$). This distinction is relevant because depending on the convention employed, similar studies could appear contradictory.

VBR studies of TLE have shown significant T_2 increases in the hippocampus, amygdala, parahippocampal gyrus, anterior temporal white matter typically ipsilateral to seizure onset [33–35]. Moreover, T_2 hyperintensities reported by VBR are more widespread in MTS

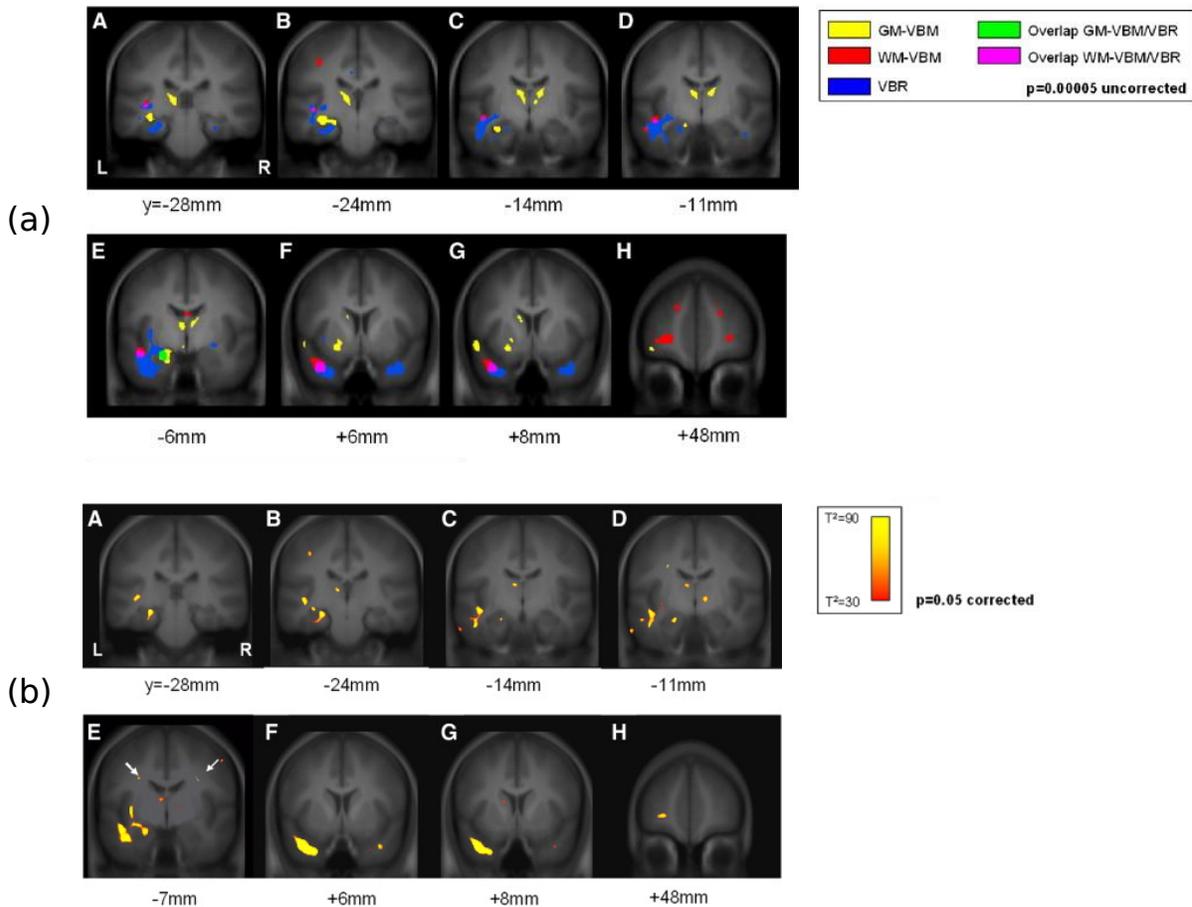


Figure 1.14: **VBM vs. VBR comparison.** a) univariate analyses including: VBM on gray matter (VBM-GM), white matter (VBM-WM), and VBR; b) Joint multivariate analysis with Hotelling's T^2 statistic. It can be seen in (a) that VBR shows widespread changes, particularly in the temporal white matter, whereas changes in gray matter are observed in the hippocampus and thalamus among other regions with VBM. Notice also, the correspondence among the regions highlighted as univariate overlaps in (a) and the results of the joint multivariate analysis in (b).¹⁵

patients compared to non-MTS patients [34]. Unlike VBM, VBR results are not tissue-specific: In VBR images are not segmented into tissue classes, and in this respect it might have an advantage over VBM, where misclassification of voxels into tissue classes (scalp to neocortex, or partial volumes in tissue boundaries) can affect subsequent statistical analysis. However, VBM and VBR results can be complimentary. Particularly, in the study of patients with mesial temporal sclerosis, VBM + VBR has produced regions where volumetric reductions are noticeable (VBM), regions with characteristic hyperintensities (VBR), and regions where both

¹⁵Adapted from NeuroImage, 39 (2008), G. Pell et al., *Composite voxel-based analysis of volume and T2 relaxometry in temporal lobe epilepsy*, 1151-1161, © 2015, with permission from Elsevier.

occur simultaneously, particularly in the hippocampus, amygdala and temporal white matter (Figure 1.14) [35].

Voxel-based statistics of DTI-derived maps (VBS)

Approaches similar to VBR have been followed to compare DTI-derived images, such as FA and MD maps. These studies are referred in the literature as voxel-based statistics (VBS), VBS-type analysis, or voxel-based DTI, to make the distinction that the images being analyzed do not describe magnetic relaxation processes. In my opinion, a more descriptive term such as VBDA (voxel-based diffusion analysis) could be employed. However, in this section I refer to these studies as VBS, for simplicity and to maintain the tradition of employing three-letter acronyms.

Similar to VBR, spatial normalization in VBS occurs by proxy. In this case, the non-diffusion weighted image (b-value=0) is used for registration to the anatomical template (or the non-diffusion weighted average image in case of having acquired more than one). Then, the calculated registration parameters are applied to the FA and MD images. After that, the smoothing step is similar to VBM and VBR, with isotropic Gaussian kernel filters between 4mm [65] and 8mm [39, 42] full-width-at-half-maximum (FWHM).

VBS studies of TLE generally show a pattern of changes characterized by FA loss and MD increase. More specifically, VBS has shown ipsilateral increase of MD in temporal and extratemporal areas including the hippocampus, parahippocampal and frontoparietal regions; and extensive temporal white matter changes in MTS patients (primarily in left TLE (L-TLE) patients) [39, 42]. By comparison, no involvement of the mesial temporal structures was identified in non-MTS patients, but rather, an increase in thalamic MD has been reported [65]. In addition, substantially decreased FA in temporal white matter has been reported, mainly ipsilaterally to the seizure onset but with some contralateral findings, and FA loss has been evidenced in the thalamus for both MTS and non-MTS groups [39, 65].

Regarding comparison of voxel-based to ROI-based techniques, it has been shown that each technique produces significant results independently but do not necessarily match. Hence, they

can be used in tandem to have a more comprehensive view of the pathological process under examination [66, 67].

1.3.3 Structural analysis of white matter

Tract-based spatial statistics (TBSS)

Tract-based spatial statistics (TBSS) [12] is a method that aims to improve the two main limitations of VBS analyses of FA: proper spatial normalization, and sensitivity to amount of spatial smoothing. On the one hand, the spatial normalization employed by VBS does not guarantee a perfect alignment of white matter fibers. For example, it has been shown that VBS studies may report residual registration misalignments as FA reductions, particularly when there are variations in ventricular sizes [68]. On the other hand, it has been shown that VBS-FA results depend strongly on the amount of smoothing [69].

Briefly, TBSS has four steps: (i) registration of all subjects' FA images to a common template using non-linear registration. At this stage perfect alignment is not expected; (ii) creation of a FA average image and extraction of a *skeleton* by removing non-maximal values, perpendicularly to the local structure. Areas of low FA and/or high inter-subject variability are removed also; (iii) projection of each subject's aligned FA image onto the skeleton; and (iv) voxel-wise statistics across subjects on the skeletonized versions of the respective FA maps. Using the same method, it is also possible to project white matter MD maps onto the skeleton for statistical analysis.

TBSS studies of temporal lobe epilepsy have shown general FA reduction in the corpus callosum, thalamus, internal/external capsule, temporooccipital and fronto temporal connections [70]. Changes in the uncinate fasciculi are more prevalent in MTS patients compared to non-MTS patients [44]. MD increases have been also been detected with TBSS in clusters of temporal white matter predominantly ipsilateral to seizure onset [39, 70].

TBSS results have been compared to those obtained from VBS analyses showing that TBSS may be more sensitive to white matter changes than VBS studies [39]. However it is important

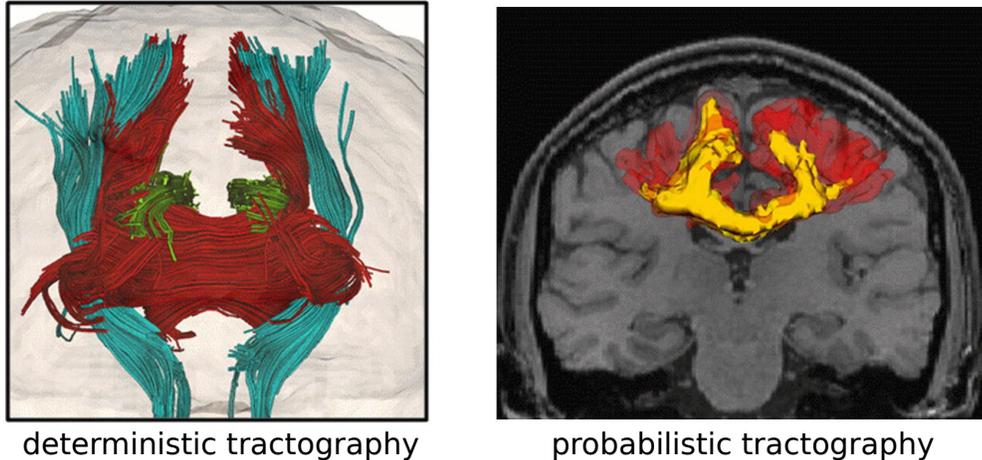


Figure 1.15: **Tractography**. Tractography constructs global map of white matter connectivity, either deterministically or probabilistically. In probabilistic tractography maps of confidence in connection to the *seeds* are obtained, here shown as isosurfaces.¹⁶

to consider the fact that the hippocampal formation does not form a part of the FA skeleton, and thus VBS could provide complimentary information for mesial structures. In the same manner, TBSS results have been combined with VBM analysis of gray matter, showing a correlation between hippocampal gray matter volume and FA white matter [44].

White matter tractography and tractometry

While MD and FA are *scalar* values characterizing the diffusion process (Figure 1.12), the eigenvectors of the diffusion tensor can be used to identify coherent anatomical patterns of diffusion, more specifically, to identify white matter fibers where diffusion is highly anisotropic. This technique is known as white matter *tractography* and allows the identification of white matter pathologies such as trauma, tumours and inflammation.

Tractography studies can use either deterministic or probabilistic methods. Deterministic tractography relies solely on information from the source DTI study, such as the main eigenvector and the FA values on every voxel to guide streamlining algorithms. In contrast, probabilistic methods use white matter atlases [71] providing a map of relative probability that a voxel belongs to a particular fiber given its location, DTI properties, and the similarity be-

¹⁵Adapted from Brain and Language 131 (2014), J. Campbell and B. Pike, *Potential and limitations of diffusion MRI tractography for the study of language*, 65-73, © 2015, with permission from Elsevier.

tween its orientation and that of the voxel's atlas [72]. Both generally require *seed voxels* to initialize the estimation of tracts. Seed selection can be manual, or semi-automatic by means of a registered pre-seeded template [43]. Small FA values in white matter can misguide the estimation of tracts. Therefore, it is common to restrict these algorithms to FA regions above a predetermined threshold (between 0.15 and 0.3) [43, 72, 73].

Tractography holds promise for identifying white matter connections *in-vivo*. Nevertheless, current techniques are limited by many factors including: (i) sensitivity to initialization parameters, including seed location; (ii) limited spatial resolution that undermines the identification of sub-voxel structures and the resolution of fiber crossing paths; (iii) the lack of a gold standard to validate results. Additionally, results often need to be manually adjusted by an expert to control for bad or misidentified tracts.

Tractometry is the process of assigning quantitative measures to the reconstructed streamlines from tractography [74]. In studies of temporal lobe epilepsy, both the average FA and MD are quantified per tract [43, 73]. Though most TLE tractometry studies agree on reporting FA loss in tracts ipsilateral to the seizure onset, a more detailed interpretation of results is challenging, since the extent and number of regions vary from study to study and these are highly dependent on the specific tracing algorithm employed. For example, two independent research groups reported that patients with L-TLE present more widespread white matter abnormalities than right TLE (R-TLE) patients [72, 75], while a third group reported the opposite [43].

Imaging analyses show pieces of a much larger puzzle in which findings are heterogeneous and highly dependent of individual patient pathologies. This poses the question: how can these clues be combined to obtain a diagnosis? Machine learning algorithms contribute to this goal by discovering distinctive patterns in clinical data that can be diagnostically useful.

1.4 Machine Learning

Machine learning (ML) is a discipline where algorithms are used to identify patterns in data. To accomplish this, datasets are divided into training and testing sets: the training set is used to learn the pattern and the testing set to validate the correctness of the identified pattern. When a pattern has been correctly identified, the algorithm can be used to make *individual predictions*.

In neuroimaging, ML methods have been applied to the identification of Alzheimer's disease [76, 77]; autism [78, 79]; schizophrenia and bipolar disorder [80, 81]; attention deficit hyperactivity disorder (ADHD) [82]; classification of fMRI brain states [83]; and epilepsy [51, 84–87]. Machine learning can assist in the *identification* and *lateralization* of TLE patients, serving a diagnostic purpose.

ML methods can be classified in four categories:

- supervised learning: the algorithm is shown both the inputs and desired outputs of the training set. The goal is to learn how to estimate the outputs given the inputs. The algorithm learns from a knowledgeable external supervisor, e.g., artificial neural networks, naive Bayes, support vector machines.
- unsupervised learning: the algorithm is presented with inputs and no outputs. Its goal is to identify hidden patterns in data. Cluster analysis, principal component analysis (PCA), hidden Markov models, and self-organizing maps belong to this category.
- reinforcement learning: the algorithm is presented with inputs, actions and rewards. Its goal is to learn how to map inputs to actions to maximize the reward, i.e., robots, chess programs, automata. [88].
- deep learning: the algorithm is presented with a large-scale complexly structured input. Its goal is to model high-level patterns in the data by learning multiple levels of data representation using semi-supervised or unsupervised methods at every level [89].

From these, supervised learning is the most commonly used in neuroimaging: the input is represented by image features (e.g., voxels, volumes, tracts, ROIs), and the output is the

subject's *class*. Classes (also called labels) are often binary: patient or control; left TLE or right TLE; MTS or non-MTS. Hence, *the goal of ML methods in neuroimaging, is to use image features to predict the subject's class*. By making a prediction for each subject in a sample, *automatic classification* can be achieved and hence, these ML algorithms are also called *classifiers*.

1.4.1 Feature selection

Not all features contribute in the same measure to obtaining a prediction. For example, a background voxel would not have any classification ability; whereas a voxel in the hippocampus may help to discriminate MTS patients from controls. Thus, feature selection consists of the automatic identification of relevant features for a given classification problem.

Curse of dimensionality

In neuroimaging, the number of features usually exceeds the number of subjects. For example, an image can have millions of voxels (which we could use as features); the number of subjects (observations of those features) is smaller by several orders of magnitude. This is a common problem in machine learning known as the *curse of dimensionality* [90], or *small-n-large-p* [91]. Consequently, without selecting informative features and discarding irrelevant, noisy features, a classifier is at risk of *overfitting*, resulting in a poor predictive performance. In this context, the concept of overfitting is similar to that of regression modeling: the more complex a regression model is, the more likely is to overfit. Similarly, a ML algorithm trying to learn a pattern made from hundreds of thousand of features, given a reduced number of observations, can very well be modeling a spurious noise pattern in the training set, rather than a pattern of interest.

Hence, feature selection in neuroimaging mitigates the effect of the curse of dimensionality, selecting good features and increasing the likelihood of obtaining a highly predictive classifier [92]. Moreover, feature selection facilitates data visualization and understanding [93], and may provide insight into the scientific question of interest [94]. Feature selection methods can

be divided in two categories: *supervised* and *unsupervised*.

Supervised feature selection

Supervised feature selection methods know the “ground truth” provided by an expert. In TLE, the ground truth corresponds to the diagnostic labels: patient or control, L-TLE or R-TLE, etc.

Supervised feature selection methods can be divided in three categories (Figure 1.16):

- **filter methods:** these methods do not have access to the classifier, but they have access to the labels. Hence, feature relevance can be established by comparing the feature with the label using statistical (t-test, ANOVA, Pearson correlation) or information-theoretic approaches (information gain, mutual information, entropy). Filter methods are fast and produce feature sets that can be used with any ML algorithm.
- **wrapper methods:** these methods have access to the classifier, and use its performance to measure the relevance of individual, or subsets of features. Thus, feature subsets are generated using a search strategy (e.g., statistical sampling, cross-validation) and the best feature subset leading to a correct classification is kept. A classifier using a wrapper method *may* perform better than a classifier using a filter method since, the features have been “tuned” for it. However, this dependency implies that features lack generality: the optimal feature set is specific to the classifier under consideration. Methods belonging to this category are recursive feature elimination (RFE) and Searchlight [95]. Due to the intensive use of the classifier as the ranking agent, wrapper methods are computationally expensive.
- **embedded methods:** in these methods, the linkage between feature selection and the classifier is stronger than in wrapper methods. In other words, feature selection makes part of the classifier. Regularized least square regression techniques such as the least absolute shrinkage and selection (LASSO), and Elastic Nets belong to this category [96, 97].

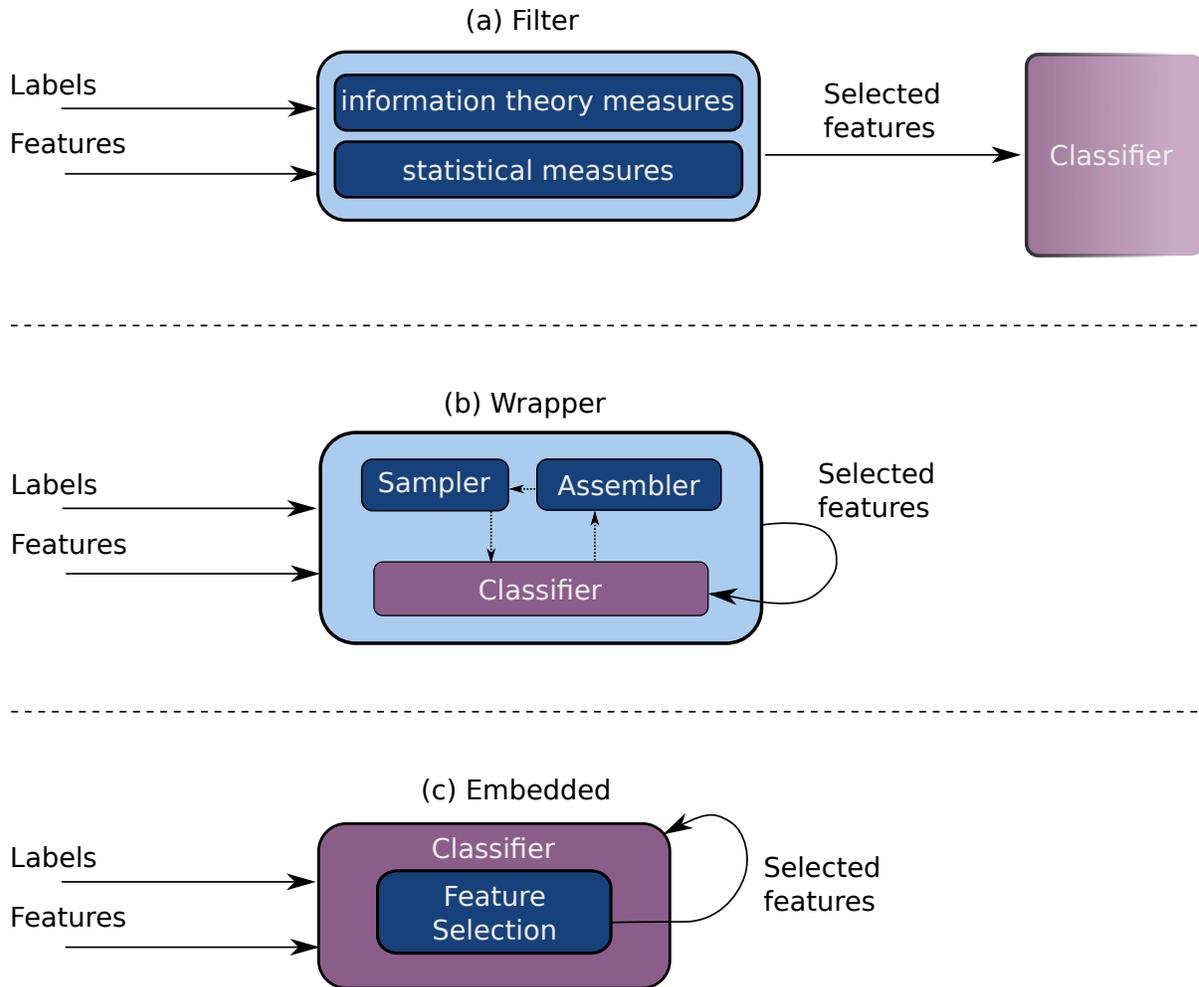


Figure 1.16: **Supervised feature selection.** a) filter methods are fast and independent from the classifier; b) wrapper methods select features according to the classification performance obtained with them; c) embedded methods are part of the classification algorithm.

Unsupervised feature selection

Unsupervised feature selection methods are oblivious to existence of a ground truth. Instead, these techniques focus on creating relevant features through linear or non-linear combinations of the original predictor variables (features) [94]. PCA, independent component analysis (ICA), clustering, spectral transformations (Fourier, Hadamard), and wavelet transforms are unsupervised feature selection methods. Hence, the main goal of unsupervised methods is to obtain an appropriate data representation in far fewer dimensions than the original feature space, thus addressing the curse of dimensionality [93].

1.4.2 Support Vector Machines

A support vector machine (SVM) is a binary classifier that computes a *decision boundary* using training data and predicts labels for new observations [98, 99].

The training dataset can be formally defined as: $\{(x_i, y_i)\}_{1 \leq i \leq N}$, $x_i \in \mathbb{R}^d$ where x_i is a multidimensional observation, y_i their respective binary label (ground truth) and N the number of elements in the training set. In neuroimaging, x_i is a vector containing features extracted from every image, y_i is the diagnostic label assigned to that image, and N the number of subjects used for training the SVM.

Compared to other classifiers, SVMs are relatively robust to the curse of dimensionality [100, 101]. SVMs use regularization avoiding, to some extent, data overfitting. Nonetheless, SVMs can still benefit from dimensionality reduction [102].

Principle

The idea behind SVMs is to find a hyperplane that maximizes the separation between two classes (Figure 1.17c). This separation is known as the *margin* and it corresponds to the distance from the decision boundary to the closest data point in the training set.

Let the hyperplane that we are looking to be defined by the standardized normal vector $\hat{\mathbf{w}}$ and intercept \mathbf{b} :

$$\hat{\mathbf{w}}^T x + \mathbf{b} = 0 \quad (1.10)$$

For any x on the plane. Let x_n be the closest point to the hyperplane, then the margin m can be defined as:

$$m = |\hat{\mathbf{w}}^T (x_n - x)| \quad (1.11)$$

This is no other than a point-plane distance, calculated by projecting $(x_n - x)$ onto the normal direction $\hat{\mathbf{w}}$ (Figure 1.17d). The absolute value here is required since a hyperplane has two normals (one pointing up, the other down). To maximize the margin, the absolute value needs to be removed. Hence, the labels $Y = \{y_i\}$ associated to the data points $X = \{x_i\}$ in

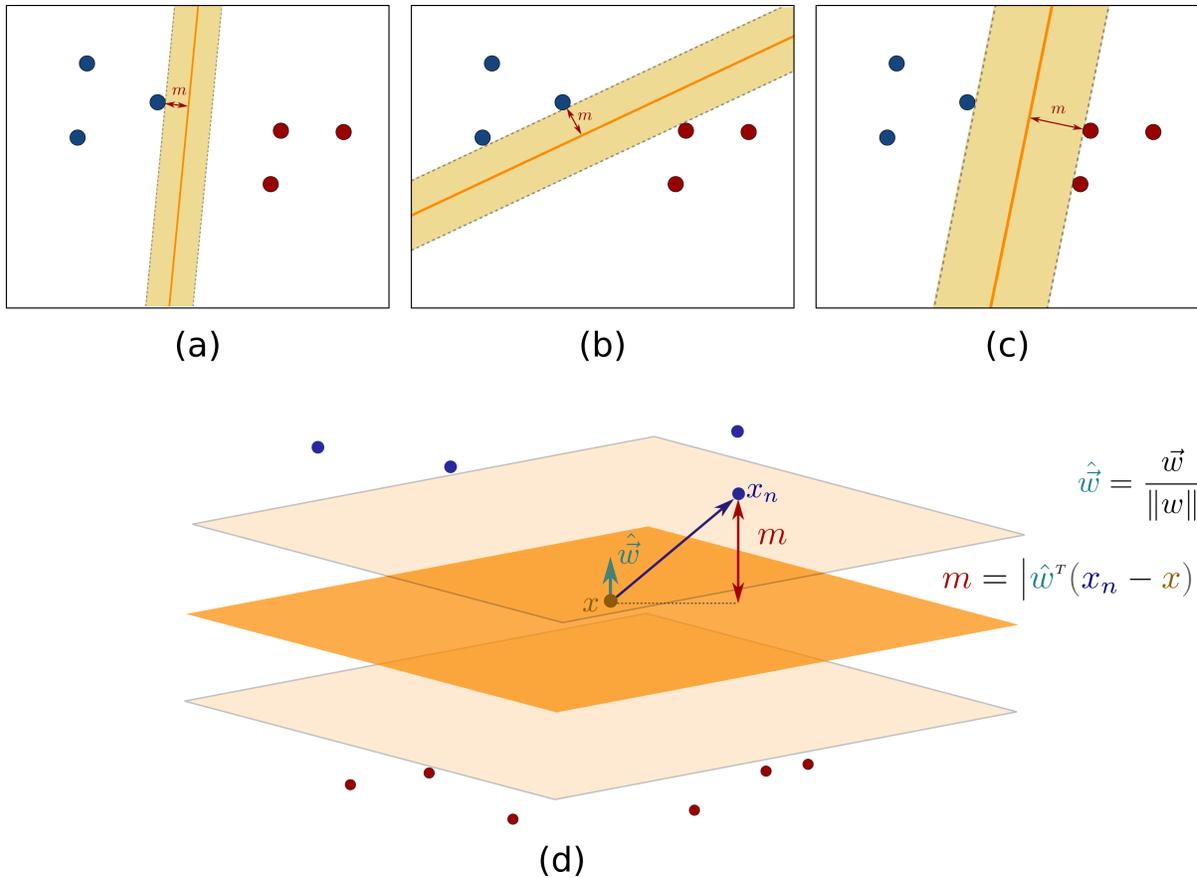


Figure 1.17: **Support vector machines.** A support vector machine (SVM) uses training data to identify the hyperplane that separates two classes maximizing the *margin*, which is defined as the distance from the hyperplane to the closest data point (d). Maximizing the margin is a quadratic optimization problem, in this figure shown in different stages from (a) to (c) where the maximum margin is obtained.

training set are replaced by $+1, -1$ according to the following rule:

$$y_i = \begin{cases} 1 & : \hat{\mathbf{w}}^T(x_i - x) > 0 \\ -1 & : \hat{\mathbf{w}}^T(x_i - x) < 0 \end{cases} \quad (1.12)$$

then the absolute value can be removed:

$$m = y_i \hat{\mathbf{w}}^T(x_n - x) \quad (1.13)$$

replacing $\hat{\mathbf{w}}$ by $\mathbf{w}/\|\mathbf{w}\|$ for *any* normal vector \mathbf{w} :

$$m = \frac{1}{\|\mathbf{w}\|} \left[y_i (\mathbf{w}^T x_n - \mathbf{w}^T x) \right] \quad (1.14)$$

but since x lies on the plane (Eq. 1.10), $\mathbf{w}^T x = 0$ and:

$$m = \frac{1}{\|\mathbf{w}\|} y_n (\mathbf{w}^T x_n) \quad (1.15)$$

without loss of generality, a restriction can be imposed so $y_n (\mathbf{w}^T x_n) = 1$. This restriction controls the scaling of the plane normal \mathbf{w} . Hence, maximizing the margin can be formulated as the optimization problem:

$$\begin{aligned} & \text{maximize} && \frac{1}{\|\mathbf{w}\|} \\ & \text{subject to} && y_i (\mathbf{w}^T x_i) \geq 1, \quad i \in [1..N] \end{aligned}$$

however, $1/\|\mathbf{w}\|$ is not easy to optimize since the vector norm includes a square root. Instead, this problem can be formulated as the equivalent minimization problem:

$$\begin{aligned} & \text{minimize} && \mathbf{w}^T \mathbf{w} \\ & \text{subject to} && y_i (\mathbf{w}^T x_i + b) \geq 1, \quad i \in [1..N] \end{aligned}$$

this formulation is known as the *primal form* and it is solvable by using Lagrange multipliers and quadratic programming. The details of the solution are beyond the scope of this thesis but it is worth mentioning that any non-zero Lagrange multipliers correspond to *support vectors*, which are data points that lie directly on the optimal margin.

Kernels and the kernel trick

In machine learning, kernels are functions that express *similarity* between data points (Table 1.1). The kernel trick consist in replacing the inner product that appears in the Lagrangian expression of the SVM optimization problem, with a kernel function κ :

$$\mathcal{L}(\beta) = \sum_{i=1}^N \beta_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N y_i y_j \beta_i \beta_j (\mathbf{x}_i \cdot \mathbf{x}_j) \quad (1.16)$$

with:

$$\mathcal{L}(\beta) = \sum_{i=1}^N \beta_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N y_i y_j \beta_i \beta_j \kappa(\mathbf{x}_i, \mathbf{x}_j) \quad (1.17)$$

Thus, similarity is not evaluated in the original data space, but rather in a space determined by the kernel function. Kernels enable the use of support vector machines with data that is otherwise not linearly separable. It is important to notice that references to non-linear support vector machines in the literature refer to *non-linear transformations of the data using kernels*. The SVM machinery is linear, always obtaining a hyperplane. Whether it is in the original space or in the transformed space, that depends on the use of kernels.

Table 1.1: **Examples of kernel functions.** Kernel functions measure similarity between two data points (here x_i and x_j). In support vector machines, kernels enable the classification of data that is not linearly separable. Here γ and r are constants, and d is the degree of the polynomial.

Type	Function $\kappa(x_i, x_j)$
linear	$x_i^T x_j$
polynomial	$\gamma (x_i^T x_j + r)^d$
Gaussian	$\exp[-\gamma \ x_i^T x_j\]$
sigmoid	$\tanh[\gamma (x_i^T x_j) + r]$

Soft margins

In most real scenarios, data are not completely linearly separable, but sufficiently close for approximation purposes. In these cases, using kernels does not guarantee an optimal outcome and may even lead to overfitting. Instead, the idea of a regularized or *soft margin* [103–105] allows obtaining a hyperplane by relaxing the constraints in the optimization problem (Equation 1.4.2) using slack factors ξ_i , and adding a penalty term governed by the regularization parameter C :

$$\begin{aligned} & \text{minimize} \quad \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^N \xi_i \\ & \text{subject to} \quad y_i (\mathbf{w}^T x_i + b) \geq 1 - \xi_i, \quad i \in \text{training examples}, \quad \xi_i \geq 0 \end{aligned}$$

Unless specified otherwise, current support vector machines in neuroimaging (including this thesis) use soft margins.

Training and testing strategies

One of the most important issues to consider when using machine learning methods is the complete separation of training and testing sets in order to avoid the problem known as *double-dipping* [92], also known as *circular analysis* in neuroimaging [95]. For example, an SVM will accurately predict each data point used in the estimation of the decision boundary, after all, the boundary is obtained by solving an optimization problem on the training data. Hence, when elements of the training set are used for testing, the *accuracy* of the classifier will be overestimated leading to the false conclusion that we have an optimal classifier. Kriegeskorte et al.(2009) befittingly explained the effect that circular analysis has in the interpretation of results [95]:

“Using the same data to train and test a linear classifier can lead us to believe that there is information about the stimulus in regions where actually there is none”

For the same reason, feature selection must not be applied to the whole dataset: by using data from the test set we are effectively biasing the selection process. This effect is also known as *overfitting by feature selection*. Hence, in an ideal scenario a dataset must be partitioned in : feature selection, training and testing sets. However, this is generally not feasible due to reduced sample sizes, exceptionally common in neuroscience and neuroimaging [106].

Popular methods in ML neuroimaging studies to avoid double-dipping are *leave-one-out* and *k-fold* cross-validation. Leave-one-out cross-validation involves selecting all but one of the subjects for training, and evaluating the classifier on the left-out subject; then, the left-out subject is reintroduced to the training set and another subject removed; this process is repeated

until all the subjects have been left-out once [83–85, 92]. Similarly, k-fold cross-validation divides the dataset in k folds of equal size; trains the SVM on k-1 folds; and then the k-th fold is used for testing; this process is repeated k times, changing each time the fold that is evaluated [51]. Similarly, bootstrapping approaches in which training and testing sets are iteratively generated by random sampling without replacement, have also been employed [76]

Finally, feature selection must be performed in the training set *only*, for the aforementioned reasons. Thus, when using cross-validation, feature selection must be performed in each resulting training set [92].

Measurements of classification performance

Since an SVM is a binary classifier, performance measures can be derived from the number of hits (correct predictions) and misses (failed predictions). For classes A and B , let A be patients (“positive” to the clinical condition), and B controls (“negative” to the clinical condition). Then, four basic measurements can be defined:

- True positives (tp): the number of hits for class A . i.e., patients correctly identified
- True negatives (tn): the number of hits for class B . i.e., controls correctly identified
- False negatives (fn): the number of misses for class A . i.e., patients identified as controls (Type II error).
- False positives (fp): the number of misses for class B . i.e., controls identified as patients (Type I error).

From this, classification performance metrics can be defined as follows:

$$\text{sensitivity} = \frac{tp}{tp + fn} \quad (1.18a)$$

$$\text{specificity} = \frac{tn}{tn + fp} \quad (1.18b)$$

$$\text{accuracy} = \frac{tp + tn}{tp + tn + fp + fn} \quad (1.18c)$$

$$\text{precision} = \frac{tp}{tp + fp} \quad (1.18d)$$

$$\text{F1-score} = \frac{2tp}{2tp + fp + fn} \quad (1.18e)$$

Sensitivity and specificity are common measures in medicine. Sensitivity measures the percentage of patients that are correctly identified; specificity, the percentage of controls that are correctly identified. Accuracy, precision and F1-score are common in pattern recognition and information retrieval. Accuracy estimates the percentage of correct classifications; precision on the other hand, the percentage of reported positives that are indeed positives; the F1-score is the harmonic mean of precision and sensitivity.

There are some scenarios where the classification occurs between patient subcategories. e.g., A = left TLE, and B = right TLE. In these cases, sensitivity and specificity lose their meaning and rather, the same formulas can be used to indicate the number of correct identifications per class for A and B respectively. The name for these measurements is replaced with “correct A” and “correct B” [84, 87]; or “sensitivity A” and “sensitivity B” [51].

1.4.3 Support vector machines in neuroimaging studies of TLE

Support vector machines have been successfully used to identify and lateralize TLE patients. For example, Keihaninejad et al.(2012) [51] used ROI features (volumes) from T1-weighted data to detect and lateralize TLE patients with MTS (accuracy 96%), and more interestingly non-MTS patients (86% – 91%). Also, perfect lateralization was obtained in MTS patients, while good lateralization was obtained in the non-MTS group (86% – 94%). Correspondingly, Focke et al.(2012) [84] identified and lateralized patients using voxel-features from T1-

weighted and DTI images, stratifying the classification per type of image (GM, WM, T2, FA, MD), and comparing L-TLE vs. controls, R-TLE vs. controls, L-TLE vs. R-TLE, and L-TLE vs. R-TLE vs. controls. The best accuracies were obtained in the gray matter (90% – 100%) and mean diffusivity (95% – 97%) images. Similarly, Zhang et al.(2012) [86] defined features from rest-fMRI to identify epilepsy patients achieving an accuracy of 83.9%, and Fang et al.(2015) [85] combined a cortical parcellation from T1-weighted images with white matter tractography to obtain a structural connectivity network. Features from this network were used to classify patients, obtaining an accuracy of 93% for lateralization, and 90% – 93.4% for patient identification.

Multivariate pattern analysis

Multi-voxel pattern analysis (MVPA) [107] describes a set of analysis techniques that applies machine learning to the detection of *local patterns* in neuroimaging. MVPA has four stages: (i) feature selection; (ii) pattern assembly; (iii) classifier training and, (iv) generalization testing. During feature selection, uninformative voxels are removed using voxel-wise statistics, or by limiting subsequent analysis to specific anatomical ROIs [108]. Another common method is to define spherical voxel-neighborhoods or “searchlights” [78, 95, 109]. Next, in the pattern assembly stage, data are sorted into discrete “brain patterns” corresponding to the difference of interest in structural MRI, or the activation pattern in fMRI. Subsequently, a classifier is trained using subset of the data, and lastly, local classifiers are tested with the rest of the pattern, hidden from training. MVPA reports classification performance for the selected ROIs.

MVPA techniques originated in the analysis of fMRI, and have been applied in several domains including: perception, learning and memory, language, intention, decision-making, emotion, anxiety, depression and dyslexia [110]. Research on MVPA is relatively contemporaneous with the writing of this document, and so far, there are not many applications on structural/DTI imaging yet, particularly in TLE. Nevertheless, a recent MVPA-fMRI study on MTS patients [108], has tested the functional capacity of the contralateral hippocampus (non-lesional side) for preserving memories, while the ipsilateral hippocampus shows no memory

preservation. Hence, MVPA has the potential to predict the effects of hippocampal resection in TLE patients that require surgery.

1.5 Thesis objectives and scope

Diagnostic clues provided by most imaging analyses are fragmented and individually insufficient to fully characterize TLE. Most techniques are highly dependent on the accuracy of the registration and the selected brain template, and even when these factors are controlled, statistical issues arise hindering inferences on individual cases. Comparative studies, in which two or more techniques are applied to the same dataset, have shown limited intersections coinciding mostly on hippocampal abnormalities and, aside from the hippocampus results are sparse. In this context, the objectives of this thesis are to:

1. Explore mechanisms for integrating MR imaging techniques into TLE diagnostic application,
2. Determine the contributions of each imaging method to TLE diagnosis,
3. Identify characteristic imaging patterns in patient subpopulations, and
4. Evaluate machine learning techniques for the identification and lateralization of TLE patients using information derived from multiple MR imaging methods.

The scope of this work corresponds to the analysis of quantitative relaxometry maps (T1, T2), DTI-derived images (FA, MD) and gray matter maps (GM) derived from VBM analyses. In regards to machine learning, classification using linear SVM and non-linear SVMs are performed.

1.5.1 Thesis overview

The rest of this document is organized in four chapters. Chapter 2 focuses on the integration of Q-MRI and DTI maps for individual patient assessment using a regions-of-interest approach.

Chapter 3 explores the quality of several SVM classification models for identifying TLE patients. Chapter 4 evaluates SVM-based classification in the temporal lobe only. Finally, chapter 5 presents conclusions and future work.

Chapter 2: Characterization of TLE through individual patient studies

In this chapter, a novel method for supplementing radiological inspection of TLE images, by providing a quantitative assessment, is presented. The method is described as follows: (i) Q-MRI and DTI-derived images are coregistered in subject space; (ii) an anatomical atlas containing cortical, white matter, and subcortical regions is registered to subject space; (iii) the ROIs defined in the atlas are used to extract features from Q-MRI and DTI-derived images for every subject; (iv) a statistical analysis is performed on these features by comparing each subject against a control group of healthy volunteers; and (v) results are reported in the form of *individual feature maps* which illustrate patient-specific findings.

Chapter 3: TLE detection using support vector machines in multi-parametric quantitative MR imaging

The regions of interest identified in Chapter 2 are used in this chapter to examine the quality of automatic patient identification and lateralization with SVMs. This chapter studies *classification models* considering these factors: (i) the type of images employed; (ii) the feature selection method; (iii) whether or not dimensionality reduction is used (PCA); and (iv) the type of SVM. Classification models are trained on TLE identification and lateralization tasks for multiple combinations of input images and performance metrics are obtained. Using these metrics, recommendations are made regarding to the optimal model for TLE analysis. Furthermore, *informative features* for each classification task are identified, with a salient prevalence of asymmetry features. Additionally, elements of machine learning are evaluated, such as the appropriateness of feature selection methods, and the stability of automatic classification in small samples.

Chapter 4: Classification of temporal lobe epilepsy patients using voxel-based asymmetry

Based on the relevance of asymmetry in the identification of patients evidenced by previous chapters, this chapter explores the classification of TLE patients using voxel-based asymmetry features. The method can be described as follows: (i) Q-MRI and DTI-derived images are registered to a common symmetric template for all subjects; (ii) T1, FA, MD, and GM asymmetry maps are generated by estimating the *local asymmetry* using spherical voxel-neighbourhoods; (iii) automatic classification experiments are performed in the anterior temporal lobe (ATL), mesial temporal region (MSL), and neocortical temporal region (NEO); and (iv) results are compared to a baseline experiment in which local asymmetry is not estimated. This chapter also presents a group-wise statistical analysis on *regional asymmetry*, defined in this case as the average of local asymmetries over the regions of analysis (ATL, MSL, NEO).

Chapter 5: Conclusions

Finally, a summary of the limitations, contributions and an outline for future research is presented.

1.6 Co-authorship Statement

Chapter 2 is adapted from the paper:

- **Diego Cantor-Rivera**, Sandrine de Ribaupierre, Jonathan C. Lau, Seyed Mirsattari, Maged Goubran, Jorge G. Burneo, David A. Steven, Terry M. Peters, Ali R. Khan, “*Individual Feature Maps: a patient-specific analysis tool with applications in temporal lobe epilepsy*”, International Journal of Computer Assisted Radiology and Surgery (submitted).

Chapter 3 is adapted from the paper:

- **Diego Cantor-Rivera**, Ali R. Khan, Maged Goubran, Seyed M. Mirsattari, Terry M. Peters, “*Detection of temporal lobe epilepsy using support vector machines in multiparametric quantitative MR imaging*”, Computerized Medical Imaging and Graphics 41 (2015) 14-28.

Chapter 4 is adapted from the paper:

- **Diego Cantor-Rivera**, Sandrine de Ribaupierre, Jonathan C. Lau, John S.H. Baxter, Seyed M. Mirsattari, Jorge G. Burneo, David A. Steven, Terry M. Peters, Ali R. Khan, “*Classification of temporal lobe epilepsy patients using voxel-based asymmetry (VBA)*”, NeuroImage: Clinical (submitted).

REFERENCES

- [1] J. F. Téllez-Zenteno, L. Hernández-Ronquillo, A review of the epidemiology of temporal lobe epilepsy, *Epilepsy research and treatment* 2012.
- [2] D. K. Angeles, Proposal for revised clinical and electroencephalographic classification of epileptic seizures, *Epilepsia* 22 (4) (1981) 489–501.
- [3] W. T. Blume, H. O. Lüders, E. Mizrahi, C. Tassinari, W. van Emde Boas, J. Engel, Glossary of descriptive terminology for ictal semiology: report of the ILAE task force on classification and terminology, *Epilepsia* 42 (9) (2001) 1212–1218.
- [4] W. T. Blume, Slowly evolving trends in temporal lobe epilepsy management at london health sciences centre, *Epilepsy Research and Treatment* 2013.
- [5] C. P. Panayiotopoulos, The new ilae report on terminology and concepts for organization of epileptic seizures: a clinicians critical view and contribution, *Epilepsia* 52 (12) (2011) 2155–2160.
- [6] A. T. Berg, S. F. Berkovic, M. J. Brodie, J. Buchhalter, J. H. Cross, W. van Emde Boas, J. Engel, J. French, T. A. Glauser, G. W. Mathern, et al., Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE commission on classification and terminology, 2005–2009, *Epilepsia* 51 (4) (2010) 676–685.
- [7] E. B. Bromfield, J. E. Cavazos, J. I. Sirven, Basic mechanisms underlying seizures and epilepsy, American Epilepsy Society, 2006.
- [8] J. W. McDonald, A. S. Fix, J. P. Tizzano, D. D. Schoepp, Seizures and brain injury in neonatal rats induced by 1s, 3r-acpd, a metabotropic glutamate receptor agonist, *The Journal of neuroscience* 13 (10) (1993) 4445–4455.
- [9] T. M. Bridges, J. M. Rook, M. J. Noetzel, R. D. Morrison, Y. Zhou, R. D. Gogliotti, P. N. Vinson, Z. Xiang, C. K. Jones, C. M. Niswender, et al., Biotransformation of a novel

positive allosteric modulator of metabotropic glutamate receptor subtype 5 contributes to seizure-like adverse events in rats involving a receptor agonism-dependent mechanism, *Drug Metabolism and Disposition* 41 (9) (2013) 1703–1714.

- [10] P. Kwan, G. J. Sills, M. J. Brodie, The mechanisms of action of commonly used antiepileptic drugs, *Pharmacology & therapeutics* 90 (1) (2001) 21–34.
- [11] E. H. Bertram, Temporal lobe epilepsy: where do the seizures really begin?, *Epilepsy & behavior* 14 (1) (2009) 32–37.
- [12] S. Smith, EEG in the diagnosis, classification, and management of patients with epilepsy, *Journal of Neurology, Neurosurgery & Psychiatry* 76 (suppl 2) (2005) ii2–ii7.
- [13] S. Raghavendra, J. Nooraine, S. M. Mirsattari, Role of electroencephalography in presurgical evaluation of temporal lobe epilepsy, *Epilepsy research and treatment* 2012.
- [14] M. Javidan, Electroencephalography in mesial temporal lobe epilepsy: a review, *Epilepsy research and treatment* 2012.
- [15] D. Silverman, R. Bernard, M. Mellies, The anterior temporal electrode and the ten-twenty system, *Electroencephalography and clinical neurophysiology* 12 (3) (1960) 735–737.
- [16] K. Oguz, Magnetic resonance imaging in epilepsy, in: N. Akalan, C. Di Rocco (Eds.), *Pediatric Epilepsy Surgery, Vol. 39 of Advances and Technical Standards in Neurosurgery*, Springer Vienna, 2012, pp. 61–83.
- [17] V. Gupta, R. Bronen A., Epilepsy, in: S. W. Atlas (Ed.), *Magnetic resonance imaging of the brain and spine, Vol. 1*, Lippincott Williams & Wilkins, 2009, pp. 307–342.
- [18] C. Limotai, S. M. Mirsattari, Role of functional MRI in presurgical evaluation of memory function in temporal lobe epilepsy, *Epilepsy research and treatment* 2012.

- [19] A. Wang, T. M. Peters, S. de Ribaupierre, S. M. Mirsattari, Functional magnetic resonance imaging for language mapping in temporal lobe epilepsy, *Epilepsy research and treatment* 2012.
- [20] H. Stefan, G. Pawlik, H. Böcher-Schwarz, H. Biersack, W. Burr, H. Penin, W.-D. Heiss, Functional and morphological abnormalities in temporal lobe epilepsy: a comparison of interictal and ictal EEG, CT, MRI, SPECT and PET, *Journal of neurology* 234 (6) (1987) 377–384.
- [21] N. Nelissen, W. V. Paesschen, K. Baete, K. V. Laere, A. Palmi, H. V. Billoen, P. Dupont, Correlations of interictal FDG-PET metabolism and ictal SPECT perfusion changes in human temporal lobe epilepsy with hippocampal sclerosis, *NeuroImage* 32 (2) (2006) 684 – 695.
- [22] A. Mansouri, A. Fallah, T. A. Valiante, Determining surgical candidacy in temporal lobe epilepsy, *Epilepsy research and treatment* 2012.
- [23] M. Honovar, B. S. Meldrum, Epilepsy, in: P. L. Lantos, D. I. Graham (Eds.), *Greenfield's Neuropathology*, Vol. 1, Arnold Ltd, 2002, pp. 899–940.
- [24] S. S. Ho, S. F. Berkovic, S. U. Berlangieri, M. R. Newton, G. F. Egan, H. J. Tochon-Danguy, W. J. McKay, Comparison of ictal SPECT and interictal PET in the presurgical evaluation of temporal lobe epilepsy, *Annals of neurology* 37 (6) (1995) 738–745.
- [25] H.-L. Margaret Cheng, N. Stikov, N. R. Ghugre, G. A. Wright, Practical medical applications of quantitative mr relaxometry, *Journal of Magnetic Resonance Imaging* 36 (4) (2012) 805–824.
- [26] D. C. Look, D. R. Locker, Time saving in measurement of nmr and epr relaxation times, *Review of Scientific Instruments* 41 (2) (1970) 250–251.
- [27] S. C. Deoni, Quantitative relaxometry of the brain, *Topics in magnetic resonance imaging: TMRI* 21 (2) (2010) 101.

- [28] K. A. Christensen, D. M. Grant, E. M. Schulman, C. Walling, Optimal determination of relaxation times of fourier transform nuclear magnetic resonance. determination of spin-lattice relaxation times in chemically polarized species, *The Journal of Physical Chemistry* 78 (19) (1974) 1971–1977.
- [29] S. C. Deoni, High-resolution T1 mapping of the brain at 3T with driven equilibrium single pulse observation of T1 with high-speed incorporation of RF field inhomogeneities (DESPOT1-HIFI), *Journal of Magnetic Resonance Imaging* 26 (4) (2007) 1106–1111.
- [30] H. Carr, Steady-state free precession in nuclear magnetic resonance, *Physical Review* 112 (5) (1958) 1693.
- [31] S. C. Deoni, B. K. Rutt, T. M. Peters, Rapid combined T1 and T2 mapping using gradient recalled acquisition in the steady state, *Magnetic Resonance in Medicine* 49 (3) (2003) 515–526.
- [32] P. Conlon, M. Trimble, D. Rogers, C. Callicott, Magnetic resonance imaging in epilepsy: a controlled study, *Epilepsy research* 2 (1) (1988) 37–43.
- [33] G. S. Pell, R. S. Briellmann, A. B. Waites, D. F. Abbott, G. D. Jackson, Voxel-based relaxometry: a new approach for analysis of T2 relaxometry changes in epilepsy, *Neuroimage* 21 (2) (2004) 707–713.
- [34] S. G. Mueller, K. D. Laxer, N. Schuff, M. W. Weiner, Voxel-based T2 relaxation rate measurements in temporal lobe epilepsy (TLE) with and without mesial temporal sclerosis, *Epilepsia* 48 (2) (2007) 220–228.
- [35] G. S. Pell, R. S. Briellmann, H. Pardoe, D. F. Abbott, G. D. Jackson, Composite voxel-based analysis of volume and T2 relaxometry in temporal lobe epilepsy, *Neuroimage* 39 (3) (2008) 1151–1161.
- [36] D. Le Bihan, C. Poupon, A. Amadon, F. Lethimonnier, Artifacts and pitfalls in diffusion MRI, *Journal of magnetic resonance imaging* 24 (3) (2006) 478–488.

- [37] M. Beg, M. I. Miller, A. Trounev, L. Younes, Computing large deformation metric mappings via geodesic flows of diffeomorphisms, *International Journal of Computer Vision* 61 (2) (2005) 139–157.
- [38] H. Huang, C. Ceritoglu, X. Li, A. Qiu, M. I. Miller, P. C. van Zijl, S. Mori, Correction of B0 susceptibility induced distortion in diffusion-weighted images using large-deformation diffeomorphic metric mapping, *Magnetic Resonance Imaging* 26 (9) (2008) 1294 – 1302.
- [39] N. K. Focke, M. Yogarajah, S. B. Bonelli, P. A. Bartlett, M. R. Symms, J. S. Duncan, Voxel-based diffusion tensor imaging in patients with mesial temporal lobe epilepsy and hippocampal sclerosis 40 (2) (2008) 728–737.
- [40] D. W. Gross, Diffusion tensor imaging in temporal lobe epilepsy., *Epilepsia* 52 Suppl 4 (2011) 32–34.
- [41] W. M. Otte, P. van Eijsden, J. W. Sander, J. S. Duncan, R. M. Dijkhuizen, K. P. J. Braun, A meta-analysis of white matter changes in temporal lobe epilepsy as studied with diffusion tensor imaging., *Epilepsia* 53 (4) (2012) 659–667.
- [42] Y.-M. Shon, Y.-I. Kim, B.-B. Koo, J.-M. Lee, H. J. Kim, W. J. Kim, K. J. Ahn, D. W. Yang, Group-specific regional white matter abnormality revealed in diffusion tensor imaging of medial temporal lobe epilepsy without hippocampal sclerosis., *Epilepsia* 51 (4) (2010) 529–535.
- [43] M. Liu, L. Concha, C. Lebel, C. Beaulieu, D. W. Gross, Mesial temporal sclerosis is linked with more widespread white matter changes in temporal lobe epilepsy, *NeuroImage: Clinical* 1 (1) (2012) 99–105.
- [44] C. Scanlon, S. G. Mueller, I. Cheong, M. Hartig, M. W. Weiner, K. D. Laxer, Grey and white matter abnormalities in temporal lobe epilepsy with and without mesial temporal sclerosis, *Journal of Neurology* 260 (9) (2013) 2320–2329.

- [45] N. Bernasconi, A. Bernasconi, Z. Caramanos, S. Antel, F. Andermann, D. Arnold, Mesial temporal damage in temporal lobe epilepsy: a volumetric MRI study of the hippocampus, amygdala and parahippocampal region, *Brain* 126 (2) (2003) 462–469.
- [46] N. Bernasconi, D. Kinay, F. Andermann, S. Antel, A. Bernasconi, Analysis of shape and positioning of the hippocampal formation: an MRI study in patients with partial epilepsy and healthy controls, *Brain* 128 (10) (2005) 2442–2452.
- [47] D. W. Gross, L. Concha, C. Beaulieu, Extratemporal white matter abnormalities in mesial temporal lobe epilepsy demonstrated with diffusion tensor imaging, *Epilepsia* 47 (8) (2006) 1360–1363.
- [48] J. W. Haller, G. E. Christensen, S. C. Joshi, J. W. Newcomer, M. I. Miller, J. G. Csernansky, M. W. Vannier, Hippocampal MR imaging morphometry by means of general pattern matching., *Radiology* 199 (3) (1996) 787–791.
- [49] J. W. Haller, A. Banerjee, G. E. Christensen, M. Gado, S. Joshi, M. I. Miller, Y. Sheline, M. W. Vannier, J. G. Csernansky, Three-dimensional hippocampal mr morphometry with high-dimensional transformation of a neuroanatomic atlas., *Radiology* 202 (2) (1997) 504–510.
- [50] S. S. Keller, J.-C. Schoene-Bake, J. S. Gerdes, B. Weber, M. Deppe, Concomitant fractional anisotropy and volumetric abnormalities in temporal lobe epilepsy: Cross-sectional evidence for progressive neurologic injury, *PloS one* 7 (10) (2012) e46791.
- [51] S. Keihaninejad, R. A. Heckemann, I. S. Gousias, J. V. Hajnal, J. S. Duncan, P. Aljabar, D. Rueckert, A. Hammers, Classification and lateralization of temporal lobe epilepsies with and without hippocampal atrophy based on whole-brain automatic MRI segmentation, *PLoS ONE* 7 (4) (2012) e33096.
- [52] A. Coan, B. Kubota, F. Bergo, B. Campos, F. Cendes, 3T MRI quantification of hippocampal volume and signal in mesial temporal lobe epilepsy improves detection of hippocampal sclerosis, *American Journal of Neuroradiology*.

- [53] N. Memarian, P. M. Thompson, J. Engel, R. J. Staba, Quantitative analysis of structural neuroimaging of mesial temporal lobe epilepsy, *Imaging in medicine* 5 (3) (2013) 219–235.
- [54] M. Esmailzadeh, H. Soltanian-Zadeh, K. Jafari-Khouzani, Mesial temporal lobe epilepsy lateralization using SPHARM-based features of hippocampus and SVM, in: *Proc. SPIE*, Vol. 8314, 2012, pp. 83144H–83144H–10.
- [55] S. S. Keller, J. O’Muircheartaigh, C. Traynor, K. Towgood, G. J. Barker, M. P. Richardson, Thalamotemporal impairment in temporal lobe epilepsy: A combined MRI analysis of structure, integrity, and connectivity, *Epilepsia* 55 (2) (2014) 306–315.
- [56] J. Ashburner, K. J. Friston, Voxel-based morphometry: the methods, *Neuroimage* 11 (6) (2000) 805–821.
- [57] S. S. Keller, N. Roberts, Voxel-based morphometry of temporal lobe epilepsy: An introduction and review of the literature, *Epilepsia* 49 (5) (2008) 741–757.
- [58] C. D. Good, I. S. Johnsrude, J. Ashburner, R. N. Henson, K. J. Friston, R. S. Frackowiak, A voxel-based morphometric study of ageing in 465 normal adult human brains, *NeuroImage* 14 (1) (2001) 21 – 36.
- [59] S. G. Mueller, K. D. Laxer, N. Cashdollar, S. Buckley, C. Paul, M. W. Weiner, Voxel-based optimized morphometry (vbm) of gray and white matter in temporal lobe epilepsy (tle) with and without mesial temporal sclerosis, *Epilepsia* 47 (5) (2006) 900–907.
- [60] S. S. Keller, M. Wilke, U. C. Wiesmann, V. A. Sluming, N. Roberts, Comparison of standard and optimized voxel-based morphometry for analysis of brain changes associated with temporal lobe epilepsy, *Neuroimage* 23 (3) (2004) 860–868.
- [61] N. K. Focke, M. R. Symms, J. L. Burdett, J. S. Duncan, Voxel-based analysis of whole brain FLAIR at 3T detects focal cortical dysplasia, *Epilepsia* 49 (5) (2008) 786–793.

- [62] P. M. House, M. Lanz, B. Holst, T. Martens, S. Stodieck, H.-J. Huppertz, Comparison of morphometric analysis based on T1-and T2-weighted MRI data for visualization of focal cortical dysplasia, *Epilepsy research* 106 (3) (2013) 403–409.
- [63] S. H. Eriksson, M. Thom, M. R. Symms, N. K. Focke, L. Martinian, S. M. Sisodiya, J. S. Duncan, Cortical neuronal loss and hippocampal sclerosis are not detected by voxel-based morphometry in individual epilepsy surgery patients, *Human brain mapping* 30 (10) (2009) 3351–3360.
- [64] S. Mehta, T. J. Grabowski, Y. Trivedi, H. Damasio, Evaluation of voxel-based morphometry for focal lesion detection in individuals, *Neuroimage* 20 (3) (2003) 1438–1454.
- [65] S. S. Keller, T. Ahrens, S. Mohammadi, J. S. Gerdes, G. Möddel, C. Kellinghaus, H. Kugel, B. Weber, E. B. Ringelstein, M. Deppe, Voxel-based statistical analysis of fractional anisotropy and mean diffusivity in patients with unilateral temporal lobe epilepsy of unknown cause, *Journal of Neuroimaging* 23 (3) (2011) 352–359.
- [66] N. R. Giuliani, V. D. Calhoun, G. D. Pearlson, A. Francis, R. W. Buchanan, Voxel-based morphometry versus region of interest: a comparison of two methods for analyzing gray matter differences in schizophrenia, *Schizophrenia research* 74 (2) (2005) 135–147.
- [67] L. Snook, C. Plewes, C. Beaulieu, Voxel based versus region of interest analysis in diffusion tensor imaging of neurodevelopment, *Neuroimage* 34 (1) (2007) 243–252.
- [68] T. J. Simon, L. Ding, J. P. Bish, D. M. McDonald-McGinn, E. H. Zackai, J. Gee, Volumetric, connective, and morphologic changes in the brains of children with chromosome 22q11. 2 deletion syndrome: an integrative study, *Neuroimage* 25 (1) (2005) 169–180.
- [69] D. K. Jones, M. R. Symms, M. Cercignani, R. J. Howard, The effect of filter size on VBM analyses of DT-MRI data, *Neuroimage* 26 (2) (2005) 546–554.
- [70] Z. Liu, Y. Xu, J. An, J. Wang, X. Yin, R. Huang, X. Lv, L. Chen, W. Wang, S. Qiu,

Altered brain white matter integrity in temporal lobe epilepsy: A tbss study, *Journal of Neuroimaging*.

- [71] S. Wakana, H. Jiang, L. M. Nagae-Poetscher, P. C. Van Zijl, S. Mori, Fiber tract-based atlas of human white matter anatomy 1, *Radiology* 230 (1) (2004) 77–87.
- [72] N. Kemmotsu, H. M. Girard, B. C. Bernhardt, L. Bonilha, J. J. Lin, E. S. Tecoma, V. J. Iragui, D. J. Hagler, E. Halgren, C. R. McDonald, MRI analysis in temporal lobe epilepsy: cortical thinning and white matter disruptions are related to side of seizure onset, *Epilepsia* 52 (12) (2011) 2257–2266.
- [73] L. Concha, C. Beaulieu, D. L. Collins, D. W. Gross, White-matter diffusion abnormalities in temporal-lobe epilepsy with and without mesial temporal sclerosis, *Journal of Neurology, Neurosurgery & Psychiatry* 80 (3) (2009) 312–319.
- [74] J. S. Campbell, G. B. Pike, Potential and limitations of diffusion MRI tractography for the study of language, *Brain and language* 131 (2014) 65–73.
- [75] M. E. Ahmadi, D. Hagler, C. R. McDonald, E. Tecoma, V. Iragui, A. M. Dale, E. Halgren, Side matters: diffusion tensor imaging tractography in left and right temporal lobe epilepsy, *American journal of neuroradiology* 30 (9) (2009) 1740–1747.
- [76] B. Magnin, L. Mesrob, S. Kinkinghunn, M. Plgrini-Issac, O. Colliot, M. Sarazin, B. Dubois, S. Lehricy, H. Benali, Support vector machine-based classification of alzheimer's disease from whole-brain anatomical MRI, *Neuroradiology* 51 (2) (2009) 73–83.
- [77] M. Dyrba, M. Ewers, M. Wegrzyn, I. Kilimann, C. Plant, A. Oswald, T. Meindl, M. Pievani, A. Bokde, A. Fellgiebel, M. Filippi, H. Hampel, S. Klöppel, K. Hauenstein, T. Kirste, S. Teipel, Combining DTI and MRI for the automated detection of alzheimer's disease using a large european multicenter dataset, in: P.-T. Yap, T. Liu, D. Shen, C.-F. Westin, L. Shen (Eds.), *Multimodal Brain Image Analysis*, Vol. 7509 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2012, pp. 18–28.

- [78] L. Q. Uddin, V. Menon, C. B. Young, S. Ryali, T. Chen, A. Khouzam, N. J. Minshew, A. Y. Hardan, Multivariate searchlight classification of structural magnetic resonance imaging in children and adolescents with autism, *Biological psychiatry* 70 (9) (2011) 833–841.
- [79] C. Ecker, V. Rocha-Rego, P. Johnston, J. Mourao-Miranda, A. Marquand, E. M. Daly, M. J. Brammer, C. Murphy, D. G. Murphy, M. A. Consortium, et al., Investigating the predictive value of whole-brain structural MR scans in autism: a pattern classification approach, *Neuroimage* 49 (1) (2010) 44–56.
- [80] H. G. Schnack, M. Nieuwenhuis, N. E. M. van Haren, L. Abramovic, T. W. Scheewe, R. M. Brouwer, H. E. Hulshoff Pol, R. S. Kahn, Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects, *NeuroImage* 84 (2014) 299–306.
- [81] J. Kambeitz, L. Kambeitz-Ilanovic, S. Leucht, S. Wood, C. Davatzikos, B. Malchow, P. Falkai, N. Koutsouleris, Detecting neuroimaging biomarkers for schizophrenia: A meta-analysis of multivariate pattern recognition studies., *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*.
- [82] R. Iannaccone, T. U. Hauser, J. Ball, D. Brandeis, S. Walitza, S. Brem, Classifying adolescent attention-deficit/hyperactivity disorder (ADHD) based on functional and structural imaging, *European Child & Adolescent Psychiatry* (2014) 1–11.
- [83] J. Mourão-Miranda, A. L. Bokde, C. Born, H. Hampel, M. Stetter, Classifying brain states and determining the discriminating activation patterns: support vector machine on functional MRI data, *NeuroImage* 28 (4) (2005) 980–995.
- [84] N. K. Focke, M. Yogarajah, M. R. Symms, O. Gruber, W. Paulus, J. S. Duncan, Automated MR image classification in temporal lobe epilepsy, *NeuroImage* 59 (1) (2012) 356 – 362.

- [85] P. Fang, J. An, L.-L. Zeng, H. Shen, F. Chen, W. Wang, S. Qiu, D. Hu, Multivariate pattern analysis reveals anatomical connectivity differences between left and right mesial temporal lobe epilepsy, *NeuroImage: Clinical*.
- [86] J. Zhang, W. Cheng, Z. Wang, Z. Zhang, W. Lu, G. Lu, J. Feng, Pattern classification of large-scale functional brain networks: Identification of informative neuroimaging markers for epilepsy, *PLoS ONE* 7 (5) (2012) e36733.
- [87] D. Cantor-Rivera, A. R. Khan, M. Goubran, S. M. Mirsattari, T. M. Peters, Detection of temporal lobe epilepsy using support vector machines in multi-parametric quantitative MR imaging, *Computerized Medical Imaging and Graphics*.
- [88] R. S. Sutton, A. G. Barto, *Introduction to reinforcement learning*, MIT Press, 1998.
- [89] G. Hinton, S. Osindero, Y.-W. Teh, A fast learning algorithm for deep belief nets, *Neural computation* 18 (7) (2006) 1527–1554.
- [90] R. Bellman, *Adaptive control processes: a guided tour*, Vol. 4, Princeton university press, 1961.
- [91] G. Fort, S. Lambert-Lacroix, Classification using partial least squares with penalized logistic regression, *Bioinformatics* 21 (7) (2005) 1104–1111.
- [92] B. Johnston, B. Mwangi, K. Matthews, D. Coghill, J. Steele, Predictive classification of individual magnetic resonance imaging scans from children and adolescents, *European child & adolescent psychiatry* 22 (12) (2013) 733–744.
- [93] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, *J. Mach. Learn. Res.* 3 (2003) 1157–1182.
- [94] B. Mwangi, T. Tian, J. Soares, A review of feature reduction techniques in neuroimaging, *Neuroinformatics* (2013) 1–16.

- [95] N. Kriegeskorte, R. Goebel, P. Bandettini, Information-based functional brain mapping, *Proceedings of the National Academy of Sciences of the United States of America* 103 (10) (2006) 3863–3868.
- [96] R. Tibshirani, Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society. Series B (Methodological)* (1996) 267–288.
- [97] H. Zou, T. Hastie, Regularization and variable selection via the elastic net, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67 (2) (2005) 301–320.
- [98] B. E. Boser, I. M. Guyon, V. N. Vapnik, A training algorithm for optimal margin classifiers, in: *Proceedings of the fifth annual workshop on Computational learning theory, COLT '92*, ACM, New York, NY, USA, 1992, pp. 144–152.
- [99] C. Cortes, V. Vapnik, Support-vector networks, *Mach. Learn.* 20 (3) (1995) 273–297.
- [100] Y. Guo, A. Graber, R. N. McBurney, R. Balasubramanian, Sample size and statistical power considerations in high-dimensionality data settings: a comparative study of classification algorithms, *BMC bioinformatics* 11 (1) (2010) 447.
- [101] J. Hua, Z. Xiong, J. Lowey, E. Suh, E. R. Dougherty, Optimal number of features as a function of sample size for various classification rules, *Bioinformatics* 21 (8) (2005) 1509–1515.
- [102] I. Guyon, J. Weston, S. Barnhill, V. Vapnik, Gene selection for cancer classification using support vector machines, *Machine learning* 46 (1-3) (2002) 389–422.
- [103] C. Cortes, V. Vapnik, Support-vector networks, *Machine learning* 20 (3) (1995) 273–297.
- [104] O. Chapelle, Training a support vector machine in the primal, *Neural Computation* 19 (5) (2007) 1155–1178.

- [105] V. D. Sánchez A, Advanced support vector machines and kernel methods, *Neurocomputing* 55 (1-2) (2003) 5–20.
- [106] K. S. Button, J. P. Ioannidis, C. Mokrysz, B. A. Nosek, J. Flint, E. S. Robinson, M. R. Munafò, Power failure: why small sample size undermines the reliability of neuroscience, *Nature Reviews Neuroscience* 14 (5) (2013) 365–376.
- [107] K. A. Norman, S. M. Polyn, G. J. Detre, J. V. Haxby, Beyond mind-reading: multi-voxel pattern analysis of fMRI data, *Trends in cognitive sciences* 10 (9) (2006) 424–430.
- [108] H. M. Bonnici, M. Sidhu, M. J. Chadwick, J. S. Duncan, E. A. Maguire, Assessing hippocampal functional reserve in temporal lobe epilepsy: a multi-voxel pattern analysis of fMRI data, *Epilepsy research* 105 (1) (2013) 140–149.
- [109] S. Janshy, O. Perez, Y. Yeshurun, T. Hendler, N. Intrator, Searchlight based feature extraction, in: *Machine Learning and Interpretation in Neuroimaging*, Springer, 2012, pp. 17–25.
- [110] Z. Yang, F. Fang, X. Weng, Recent developments in multivariate pattern analysis for functional MRI, *Neuroscience bulletin* 28 (4) (2012) 399–408.

Chapter 2

Characterization of TLE through individual patient studies

2.1 Introduction

Identifying MRI abnormalities that can be correlated with seizure activity is fundamental for the lateralization (left or right hemisphere seizure onset), seizure localization, surgical planning, and prognosis of patients with temporal lobe epilepsy (TLE). In some cases, the lack of tissue contrast and/or the abnormalities reduced size can make diagnosis difficult. In these scenarios further tests are required, generating delays and additional costs in patient care.

Computer-assisted imaging tools have improved detection and quantification of MR image abnormalities in TLE patients. Current techniques analyze changes in gray and white matter volumes using T1- and T2-weighted sequences [1, 2]; cortical thickness and cortical folding [3, 4]; hippocampal intensity, shape and volume [5–8]; as well as changes in the structure of white matter using diffusion tensor imaging (DTI), reporting fractional anisotropy (FA) and mean diffusivity (MD) [6, 9].

Many of these computer-assisted analysis methods rely on group comparisons: a TLE group is compared against a healthy control group and statistically significant differences are extrapolated to the patient population. For example, ipsilateral reduction of FA, in most cases

accompanied by increased ipsilateral MD, has been found in TLE patients through this approach [10–12]. However, group effects are not relevant for the individual presurgical assessment of TLE patients, in which the role of MRI is to detect focal abnormalities unique to each patient [2]. The heterogeneous etiology of TLE, as well as, the diverse presentation of epileptic foci, require that patient-specific techniques enable the identification of focal changes even if these are not shared by the particular patient population under study.

In this context, a novel, multi-parametric approach to patient-specific analysis integrating modern MR techniques is presented. More specifically quantitative relaxometry (T1 and T2 maps) and diffusion tensor imaging (fractional anisotropy and mean diffusivity) are analyzed, providing a comprehensive view of changes that otherwise would not be available by simple visual inspection. This approach uses focal intensity and volumetric differences between left and right hemispheres, providing a unique characterization, as well as a laterality score for every patient in the study. Although there are some limitations, the proposed technique holds promise for providing new insights into the nature, extent and anatomical relationships of structures involved in patient-specific pathologies, and might lead to new approaches in the clinical evaluation and surgical planning for TLE patients.

2.2 Method

2.2.1 Participants and Imaging Protocol

A total of 32 individuals participated in this study, 17 of whom were control volunteers (age 32 ± 10 , 12 male, 7 female) and 15 TLE patients (age 35 ± 10 , 7 male, 8 female) who were scheduled for temporal lobe surgery (8 left, 7 right). Only patients with lateralizable seizures (confirmed by prolonged video-EEG) were selected for this study. A clinical summary of our patient cohort is presented in Table 2.1. Preoperative anatomical MRI identified mesial temporal sclerosis in 8 patients (MTS+) [#1-8] (confirmed in 7 of them by post-op pathology). The remaining 7 patients [#9-15] were classified as negative for MTS (MTS-) with 3 of them having normal MR scans [#13-15].

Table 2.1: **Clinical summary.** Hipp: hippocampus; DNET: dysembrioplastic neuroepithelial tumour; CD: cortical dysplasia; FCD: Focal cortical dysplasia; LATL: left anterior temporal lobectomy, RATL: right anterior temporal lobectomy. (+) indicates EEG findings outside the temporal lobe

Patient	Group (based on MRI)	Age	Gender	E. Dur. (years)	Risk factors for Epilepsy	EEG (spikes and seizures)	Anatomical MRI (1.5T)	Type of Epilepsy	Surgery	Pathology	Engel outcome	Follow up (months)
1	L-MTS	20	M	17	Varicella Zoster encephalitis with status epilepticus at age 3 months.	Left temporal and occipital spikes; left temporal seizures	Left MTS; encephalomalatia of the left occipital lobe	Lesional LTLE+	LATL	MTS	II	19
2	L-MTS	48	F	12	Post cardiac transplant	Bitemporal spikes (maximum left) and left temporal seizures	Bilateral MTS (maximum left)	Bitemporal TLE, maximum left	LATL	MTS	II	21
3	L-MTS	30	M	28	no	Left temporal spikes and seizures	Left MTS; possible left temporal cortical dysplasia; poorly developed Left temporal pole white matter	Lesional LTLE	LATL	MTS	I	21
4	L-MTS	32	F	13	no	Left temporal spikes and seizures	Left MTS	Lesional LTLE	LATL	MTS	I	15
5	L-MTS	34	M	24	Febrile seizures at age 10 months	Left temporal, parietal and central, frontal spikes, Left temporal seizures	Left MTS	Lesional LTLE+	LATL	MTS	III	11
6	Possible R-MTS	18	M	4	no	Right temporal spikes and seizures; generalized spikes and waves	Possible right MTS; lesser myelination in right temporal lobe	Lesional RTLE+	RATL	Gliosis	I	25
7	R-MTS	43	F	40	Febrile seizures at ages 3-6 years; one febrile status epilepticus	Right temporal, left hemispheric and generalized spikes; Right temporal seizures	Right MTS	Lesional RTLE +	RATL	MTS; Gliosis	I	17
8	R-MTS	33	F	13	Febrile seizures at age 5 years	Right temporal frontal spikes; right temporal seizures	Right MTS; enlarged lateral ventricles; periventricular leukomalacia	Lesional RTLE +	RATL	MTS; Gliosis	I	11
9	MTS -	40	F	24	no	Left temporal spikes and seizures	left hemispheric convexity subdural hematoma caused by subdural grid	Non-lesional LTLE	LATL	Gliosis	I	19
10	MTS -	25	F	8	Left temporal lobe low grade astrocytoma	Left temporal spikes and seizures	Encephalomalatia from previous resection of the left temporal tumour	Lesional LTLE	LATL	Gliosis	III	18
11	MTS -	56	F	41	no	Right temporal spikes and seizures	Moderate generalized atrophy; ischemic microangiopathy; lacunar infarcts in the right caudate and bilateral cerebellum	Non-lesional RTLE	RATL	Gliosis	II	6
12	MTS -	39	M	6	no	Bitemporal spikes; right temporal seizures	Right temporal FCD; possible thickening of the left fusiform, left parahippocampus and left medial temporal pole Left hippocampus slightly increased T2 and abnormal orientation	Lesional RTLE	RATL	FCD; Gliosis	I*	5
13	MTS -	23	M	5	no	Left temporal frontal spikes; left temporal seizures	Normal	Non-lesional LTLE+	LATL	FCD	II	12
14	MTS -	41	F	6	no	Bitemporal spikes; right temporal seizures	Normal	Non-lesional RTLE	RATL	Gliosis	II	11
15	MTS -	31	M	3	no	Bitemporal spikes; right temporal seizures	Normal	Non-lesional LTLE	RATL	Gliosis	I	17

+ with EEG findings outside the temporal lobe, * few auras

For this study, subjects were scanned using a 3T MR scanner (GE Discovery MR750) with whole brain, DESPOT1-HIFI and DESPOT2-FM sequences (1mm isotropic voxel) [13–15]. Two SPGR images were acquired (flip angles of 4° and 18°) along with an inversion-prepared spoiled gradient recalled acquisition in the steady state (SPGR) to calculate a quantitative T1 map. Five balance steady state free-precession (bSSFP) images were acquired with phase cycling (flip angles of 15° , 35° , 60°) to estimate a quantitative T2 map using the DESPOT2-FM procedure [16]. A DTI sequence was also acquired with the following parameters: 2.4 mm isotropic voxels, 41 directions, b-value =1000, 4 non-weighted (b=0) volumes. This protocol was approved by the Office of Research Ethics at Western University (Canada) ¹.

2.2.2 Image Preprocessing

The images acquired with the DESPOT1-HIFI and DESPOT2-FM techniques were co-registered to the 18° flip angle SPGR image using the FLIRT registration tool from FSL² [17] to account for motion between scans prior to the computation of the T1 and T2 maps.

DTI images were processed as follows: (1) non-linear geometric distortions were corrected by deformable registration of the average of the b=0 volumes to the undistorted T1 map, using a diffeomorphic registration method [18, 19]; (2) Eddy current correction and diffusion tensor estimation was performed using FSL's diffusion toolbox FDT; (3) maps of fractional anisotropy (FA) and mean diffusivity (MD) were transformed and re-sampled to the coordinate system defined by the 1mm isotropic T1 map. This preprocessing stage yielded four co-registered quantitative maps (T1, T2, FA, MD) for each subject.

2.2.3 Segmentation of regions of interest (ROI)

For each subject, a synthetic T1-weighted image, with inherent bias-field correction, was generated from the T1 map [20]. This synthetic image was used to perform an atlas-based volu-

¹The ethics approval document is available in Appendix B

²FSL software: <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>

metric ROI segmentation using Freesurfer.³ This procedure included: (1) removal of non-brain tissue using a hybrid watershed/surface deformation procedure; (2) automated Talairach transformation⁴; (3) segmentation of the white matter, gray matter and specific structures, including hippocampus, amygdala, caudate, putamen, thalamus and ventricles; (4) The gray matter and white matter were further subdivided into regions based on gyral and sulcal architecture using a surface-based registration approach. This segmentation procedure created a set of *paired regions* between the left and right hemispheres (i.e. left and right hippocampus). A total of 35 cortical, 34 white matter and 9 subcortical ROI per hemisphere were extracted for every subject.

Both temporal and extratemporal lobe regions are included in this analysis under the consideration that TLE seizures can propagate outside of the temporal lobe, therefore affecting other brain areas (Please refer to extratemporal EEG activity in Table 2.1).

2.2.4 Feature quantification

Features in the quantitative maps (T1, T2, FA, MD) were extracted using the regions of interest obtained in the previous step.

Three types of features were extracted from each map: mean intensity, intensity asymmetry, and volume asymmetry.

- The *mean intensity* was calculated by averaging the voxel intensities in every ROI.
- The *intensity asymmetry* was measured by comparing the intensity distributions in the left-right ROI pairs with a two-sample Kolmogorov-Smirnov test (KS).
- The *volume asymmetry* was quantified as the absolute difference between corresponding left-right ROIs, normalized by the total intracranial volume.

³Freesurfer software: <http://surfer.nmr.mgh.harvard.edu/>

⁴According to the FreeSurfer documentation, this step is an affine transform to the MNI305 atlas. See <http://freesurfer.net/fswiki/talairach>

Measuring asymmetry in ROI intensities

The KS test calculates the empirical cumulative distribution function (ECDF) of two distributions and then determines the maximum distance between them (Figure 2.1). Because the vertical axis corresponds to cumulative probability, the result of the test is always a positive number in the $[0,1]$ range, which is an ideal continuous scale to express asymmetry (i.e. 0 = completely symmetric regions, 1 = completely asymmetric regions).

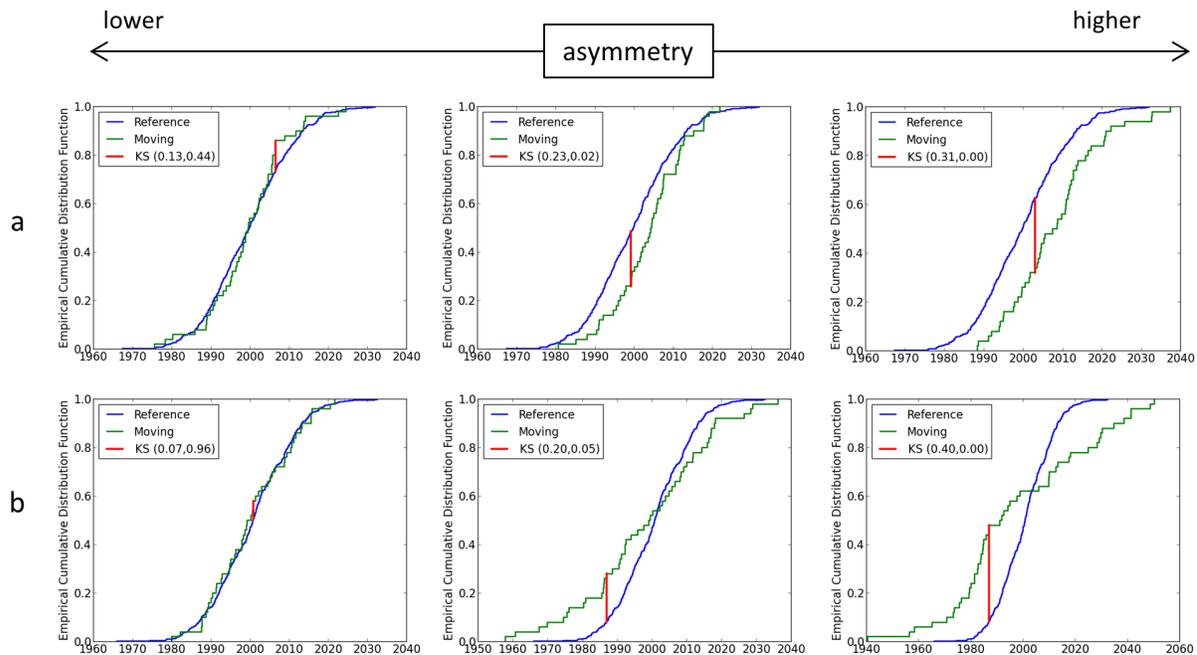


Figure 2.1: **Kolmogorov-Smirnov (KS) test**. This example shows, from left to right, increasing differences in (a) mean and (b) standard deviation of two simulated ROIs. The KS test result is shown by the vertical bar indicating the maximum distance between the left and right empirical cumulative distribution functions (ECDFs). The x-axis is an arbitrarily selected intensity range for this simulation.

There is more overlap among similar left-right ECDFs and the resulting *asymmetry* measure is small. In contrast, in the presence of a *focal pathology* the intensity distribution in one of the two ROIs can be significantly different than in the other, therefore the distance between ECDFs is large and high asymmetry is obtained. This is shown in Figure 2.2 where hippocampal MD asymmetry is measured in (a) a control subject, and (b) a TLE patient. For the control subject, the left and right hippocampal intensity distributions have a larger overlap. This is reflected in a low asymmetry value (KS = 0.17). In contrast, the same distributions have a lower overlap

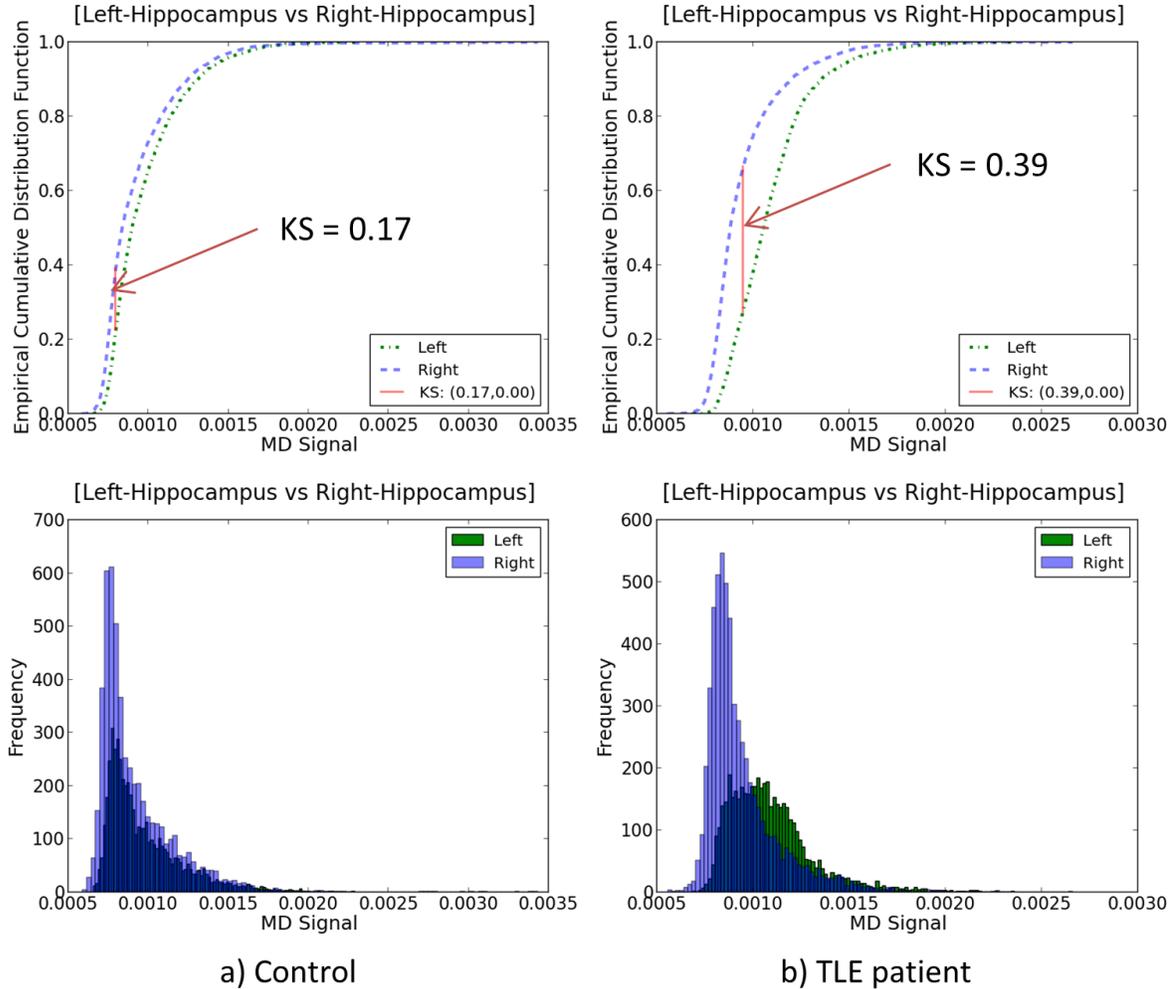


Figure 2.2: **Intensity distribution asymmetry.** A two-sample Kolmogorov-Smirnov test is employed here to compare the left and right hippocampus on MD; *intensity distribution asymmetry* is measured for **a**: a control subject and **b**: a TLE patient; The ECDFs are shown on the first row and the respective distributions, in the second row.

for the TLE patient($KS = 0.39$).

KS is a non-parametric test, which makes it ideal for dealing with TLE images where tissue intensities may depart from normality due to the presence of pathologies. Moreover, the KS test was chosen amongst the many possible measures of the difference between two probability distributions because of its robustness to small sample size considering the limited number of voxels being compared in DTI-based volumes. This robustness is a product of the tests use of the cumulative probability function which can be estimated more readily than the probability density function required by the Bhattacharyya or Kullback-Liebler divergence. These density

functions, and therefore the difference metrics between them, are heavily dependent on the bin size or selection of kernel in smoothing approaches which could lead to lower robustness in the estimation of asymmetry.

Similarly to the mean intensity and volume asymmetry features, intensity asymmetry in the control group was measured to obtain a baseline for statistical comparisons. An example is presented in Figure 2.3.

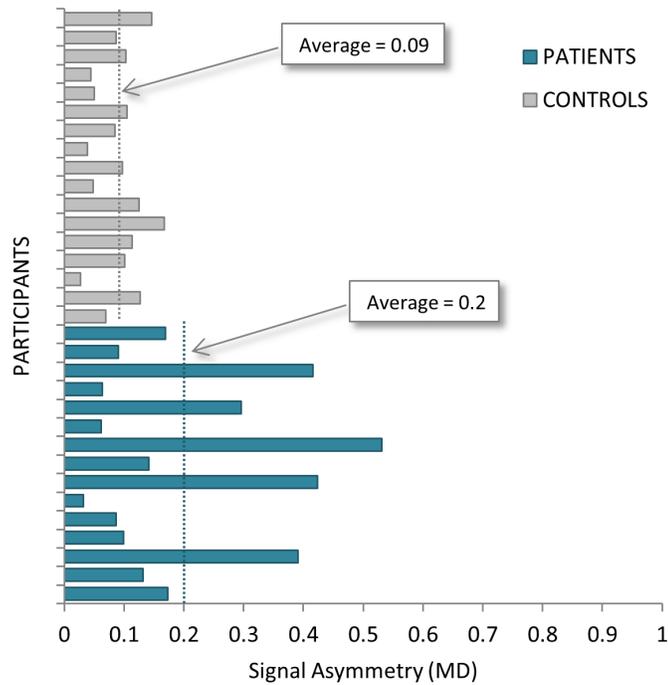


Figure 2.3: **Measuring intensity asymmetry.** Hippocampal MD asymmetry measured for patients and controls. Using the control group asymmetry as a baseline it is possible to estimate *abnormal asymmetries* in TLE patients.

2.2.5 Single-subject statistical tests

To determine the *rarity* of a feature in a given patient, single-subject t-tests [21–24] were performed. These tests compared patient features to the respective statistical distribution for the respective features obtained from the control group (Figure 2.4).

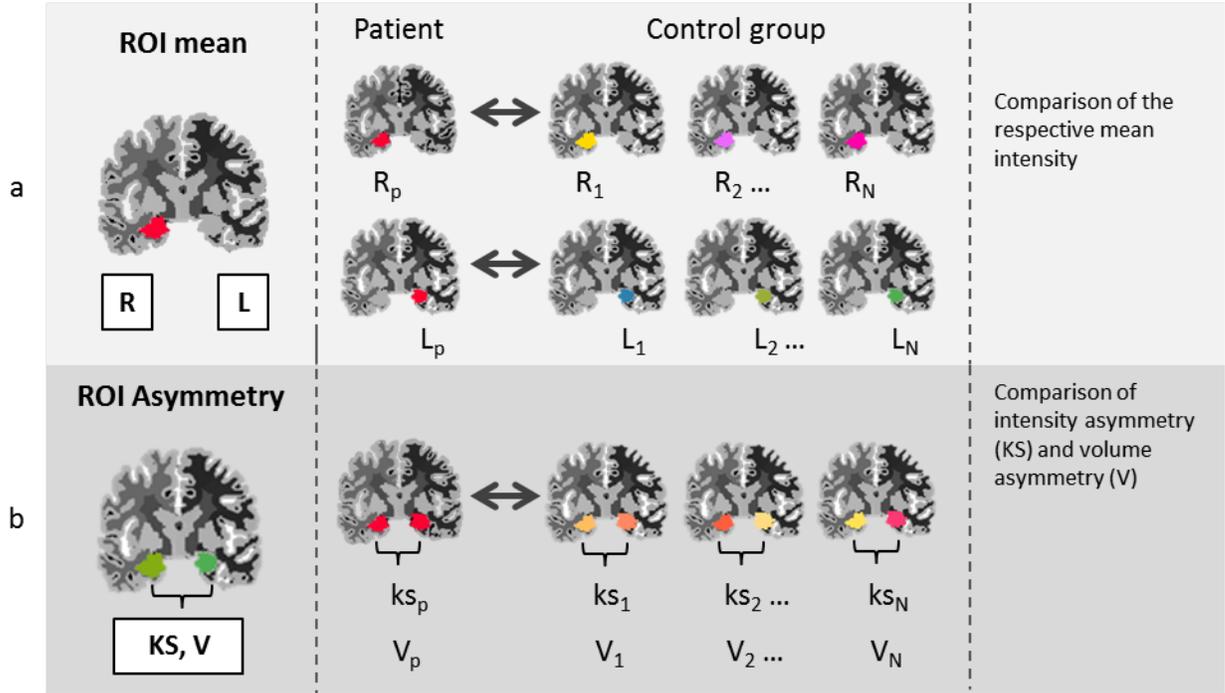


Figure 2.4: **Statistical tests.** A single-subject t-test is performed to detect individual, local hypo/hyperintensities, and abnormal asymmetry in intensity and volume. (a) mean intensity tests. L: left mean intensity, R: right mean intensity. (b) asymmetry tests. KS: intensity asymmetry, V: volume asymmetry.

Mathematically, the mean and variance of the feature in the control group are treated as population parameters and the patient's feature is treated as a sample of size one ($N=1$) not contributing to variance estimation. For a given subject s , control group c , ROI (or region pair) r , and feature X_r , this test is defined as:

$$t_{(s,X_r)} = \frac{X_r(s) - \overline{X_r(c)}}{\sigma_c \sqrt{\frac{N_c+1}{N_c}}} \quad (2.1)$$

where $X_r(s)$ is the feature X_r evaluated for subject s ; $\overline{X_r(c)}$ is the feature mean in the control group; σ_c is the feature standard deviation; and N_c is the control group size. A significance

level $\alpha = 0.05$ was employed, adjusting for multiple comparisons using Bonferroni correction.

2.2.6 Individual Feature Maps

The results of single-subject t-tests were organized in four *individual feature maps* (IFMs) to facilitate their interpretation.

- IFM{L}: this map shows the regions in the left hemisphere that are abnormally hypo/hyperintense when compared to a control population; i.e., where the ROI *mean intensity* survived statistical testing. An associated report indicates the type of finding (hypo/hyperintensity).
- IFM{R}: similar to the left map, it shows the regions in the right hemisphere that are abnormally hypo/hyperintense when compared to a control population. Information regarding the type of finding is also included in the associated report.
- IFM{KS}: shows the regions that survived the statistical testing of their *intensity asymmetry* feature. For T1, T2 and MD only the ROI on the hyperintense side is shown. For FA, only the ROI on the hypointense side is shown. This convention aids in the interpretation of the results as it has been shown by previous studies that intensity changes associated to TLE manifest as increased T1, T2 [25, 26] and MD [12] and focal loss in FA [10].
- IFM{V}: region pairs with abnormal *volume asymmetry* are reported here, however, only the smaller ROI of the pair is shown. This convention is useful for identifying relative volume reduction which is of clinical interest and it has been documented in similar TLE studies [6, 27]

A *quantitative report* accompanying these maps organizes the results according to type of IFM (L, R, KS, V), image (T1, T2, FA, MD) and type of finding: hypo or hyperintensity for

mean intensity features; (Left>Right or Right>Left) for intensity asymmetry, and (VL) smaller left volume, (VR) smaller right volume for volume asymmetry⁵.

2.2.7 Patient lateralization based on temporal lobe asymmetry

A *laterality score* to discriminate between left and right TLE patients was defined as follows: Let p be a temporal lobe ROI pair present in the {KS} map; p_l , the mean intensity in the left ROI; p_r , the mean intensity in the right ROI; and $i(p)$ be the image where the intensity asymmetry is found. The *laterality* of p , $L(p)$ is defined as:

$$L(p) = \begin{cases} -1 & \text{if } i(p) \in [T1, T2, MD] \text{ and } p_l > p_r \\ +1 & \text{if } i(p) \in [T1, T2, MD] \text{ and } p_r > p_l \\ -1 & \text{if } i(p) \in [FA] \text{ and } p_r > p_l \\ +1 & \text{if } i(p) \in [FA] \text{ and } p_l > p_r \end{cases} \quad (2.2)$$

Then, the *laterality score* S for a given patient is defined as:

$$S = \frac{\sum_p L(p)}{N} \quad (2.3)$$

where the sum extends over all the temporal lobe region pairs identified in the {KS} map, and N is the total number of pairs. The division by N normalizes the laterality score from -1 (perfect left) to +1 (perfect right).

2.2.8 Group analysis

With the IFMs computed for all the subjects in the study (including controls), an *a posteriori* statistical analysis was performed on the number of abnormalities reported by the IFMs, with the goal of identifying frequent abnormalities across the individual cases as well as to characterize different patients subgroups (L-TLE vs R-TLE, controls vs. MTS- vs. MTS+).

⁵Examples of these reports are available in Appendix A

2.3 Results

Table 2.2 shows individual results grouped by IFM, lobe, type of tissue (cortex, white matter) or other structures (hippocampus, amygdala, thalamus, caudate, ventricles, putamen, pallidum, insula and cerebellum). Statistically significant mean intensities are reported in square brackets ([R] = right, [L] = left) with the sign indicating a hyperintensity ([R+],[L+]) or a hypointensity ([R-],[L-]); intensity asymmetries are reported using parenthesis as (L>R) or (R>L); and for volume asymmetries the lesser volume is shown as (VL) for left, and (VR) for right.

In regards of laterality, 12 out of 15 patients (80%) were correctly categorized as L-TLE or R-TLE. Table 2.2 show the correspondence between the predicted laterality score and the surgery performed. The remaining three patients had a normal MRI and the study did not reveal any significant temporal intensity asymmetry (#13-15 in Table 2.2).

Hippocampal intensity abnormalities were found in all patients diagnosed with MTS (confirmed histologically), with most of these differences being detected on the MD image. Ipsilateral hippocampal volume reduction was also found also in all of them. Though patient #6 was diagnosed as a possible MTS case by the MRI, only gliosis was present in the surgical specimen. Our method did not detect hippocampal involvement in either intensity or volume for that patient.

Table 2.2: **IFM results.** Good correspondence between predicted laterality score and surgery performed; **0#**: not enough asymmetry to determine side. (L>R) intensity asymmetry with left ROI ζ right ROI, (R>L) opposite case; [L/R+]: left/right hyperintensity, [L/R-]: left/right hypointensity; Volume asymmetry is reported indicating the lesser volume as (VL) for left, and (VR) for right; ctx: cortex; wm: white matter; temp: temporal; occ: occipital; p: parietal; f: frontal;

Patient	MRI group	Left Hemisphere Map {L}	Right Hemisphere Map {R}	Signal Asymmetry Map {KS}	Volume Reduction Map {V}	Laterality Index	Surgery
1	L-MTS	<ul style="list-style-type: none"> Temporal white matter [T1+, FA-, MD+] Occ. white matter [T1+, FA-, MD+] 	<i>no findings</i>	L>R <ul style="list-style-type: none"> Temporal ctx and wm [T1, T2, MD] Parietal wm [T1], frontal wm [T2] Occipital ctx and wm[MD] Hippocampus [MD] Lateral Ventricle [T1, T2] R>L <ul style="list-style-type: none"> Temporal ctx and wm [FA] Occipital wm[FA], parietal wm[T1] 	VL <ul style="list-style-type: none"> Temporal cortex and white matter Hippocampus, putamen 	-1.00	LATL
2	L-MTS	<ul style="list-style-type: none"> Frontal white matter [T1+] Parietal white matter [T2+, FA-] Temporal white matter [FA-, MD+] Hippocampus, amygdala, thalamus, insula [MD+] 	<ul style="list-style-type: none"> Frontal white matter [T1+, FA-, MD+] Hippocampus [MD+] 	L>R <ul style="list-style-type: none"> Parietal cortex [T2] Temporal cortex [MD] 	VL <ul style="list-style-type: none"> Hippocampus 	-1.00	LATL
3	L-MTS	<ul style="list-style-type: none"> Accumbens area [FA+] 	<ul style="list-style-type: none"> Hippocampus, amygdala, thalamus, putamen, acc. area [FA+] Occipital cortex [FA-] Pallidum [FA+, MD-] 	L>R <ul style="list-style-type: none"> Hippocampus [MD] R>L <ul style="list-style-type: none"> Parietal ctx, frontal ctx, putamen, pallidum [FA] 	VL <ul style="list-style-type: none"> Hippocampus 	-1.00	LATL
4	L-MTS	<ul style="list-style-type: none"> Hippocampus [MD+] 	<i>no findings</i>	L>R <ul style="list-style-type: none"> Parahippocampal white matter [T1] Hippocampus[T2, MD] 	VL <ul style="list-style-type: none"> Hippocampus VR Frontal cortex 	-1.00	LATL
5	L-MTS	<i>no findings</i>	<ul style="list-style-type: none"> Thalamus, putamen, pallidum [FA+] 	L>R <ul style="list-style-type: none"> Amygdala, hippocampus [MD] R>L <ul style="list-style-type: none"> Amygdala, putamen, parietal cortex [FA] 	VL <ul style="list-style-type: none"> Hippocampus, Thalamus, Cuneus 	-1.00	LATL
6	Possible R-MTS	<ul style="list-style-type: none"> Entorhinal white matter[T2+] Frontal cortex (pars orbitalis) [FA+] 	<i>no findings</i>	L>R <ul style="list-style-type: none"> Accumbens area [T2] Temporal white matter [FA] R>L <ul style="list-style-type: none"> Temporal wm[T1], temporal ctx [MD] Occipital cortex [T1, T2] 	<i>no findings</i>	0.60	RATL
7	R-MTS	<i>no findings</i>	<i>no findings</i>	L>R <ul style="list-style-type: none"> Accumbens area, Caudate [T2] R>L <ul style="list-style-type: none"> Thalamus, Frontal cortex [T1], parietal cortex [FA] Hippocampus, acc. area [MD] 	VR <ul style="list-style-type: none"> Hippocampus, Thalamus 	0.20	RATL
8	R-MTS	<ul style="list-style-type: none"> Occ. wm [T1+, T2+] and occ. cortex [T2+] Frontal white matter [T1+, T2+] Parietal wm [T2+, MD+] and parietal cortex [T2+] Hippocampus [T2+], parahippocampal wm [T1+, T2+] Thalamus [T2+, MD+] Temporal cortex [T2+] Caudate [T2+], Cerebellum [T2+], inferior lateral ventricle[T1] 	<ul style="list-style-type: none"> Occ. white matter [T1+, T2+, MD+] Temp. white matter [T1+, T2+, MD+], temp. cortex [T2+, MD+] Frontal white matter [T1+, FA-, MD+], frontal cortex [T2+] Parietal wm [T1+, T2+,FA-,MD+], parietal cortex [T2+] Thalamus [T1+,T2+, FA-, MD+] Amygdala, Hippocampus, Cerebellum, Insula, Caudate [T2+] 	L>R <ul style="list-style-type: none"> Temporal cortex [T1], Thalamus [FA], Cerebellum wm [MD] R>L <ul style="list-style-type: none"> Temp. wm [T1], Frontal wm [T1,T2], LV [T1] Hippocampus [T1] Thalamus [T1] Cerebellum [FA] 	VL <ul style="list-style-type: none"> Thalamus VR Hippocampus 	0.40	RATL
9	MTS -	<i>no findings</i>	<i>no findings</i>	L>R <ul style="list-style-type: none"> Temporal white matter [T1], temporal cortex [MD], Parietal cortex [T2] R>L <ul style="list-style-type: none"> Temporal white matter [FA] 	VR <ul style="list-style-type: none"> Inferior Lateral Ventricle 	-1.00	LATL
10	MTS -	<ul style="list-style-type: none"> Temporal white matter [T1+, T2+, MD+] Temporal cortex [MD+] Frontal white matter [T2+] 	<i>no findings</i>	L>R <ul style="list-style-type: none"> Frontal cortex [T2], amygdala [FA] Temporal white matter, temp. cortex[MD] R>L <ul style="list-style-type: none"> Accumbens area [FA], frontal cortex [MD] 	VL <ul style="list-style-type: none"> Temporal cortex 	-0.67	LATL
11	MTS -	<ul style="list-style-type: none"> Fusiform cortex [MD+] 	<ul style="list-style-type: none"> Cuneus [T2+] Temporal white matter [MD+] Fusiform cortex [MD+] 	R>L <ul style="list-style-type: none"> Parahippocampal white matter [T1] Entorhinal cortex [T2] 	<i>no findings</i>	1.00	RATL
12	MTS -	<ul style="list-style-type: none"> Caudate [T2+] Insula [MD+] 	<ul style="list-style-type: none"> Frontal cortex [T2+] Temporal cortex [T2+, MD+] 	R>L <ul style="list-style-type: none"> Parahippocampal cortex [T1] Parahippocampal white matter [T1] Cerebellum white matter [T1] 	VL <ul style="list-style-type: none"> Hippocampus VR Frontal cortex 	1.00	RATL
13	MTS -	<i>no findings</i>	<ul style="list-style-type: none"> Fusiform cortex [MD+] 	<i>no findings</i>	<i>no findings</i>	0#	LATL
14	MTS -	<i>no findings</i>	<i>no findings</i>	<i>no findings</i>	<i>no findings</i>	0#	RATL
15	MTS -	<i>no findings</i>	<i>no findings</i>	<i>no findings</i>	<i>no findings</i>	0#	RATL

2.3.1 Individual feature maps for a MTS patient

This patient was a 32 year-old, right-handed woman with a 13 year history of partial seizures (Patient #4 in Table 2.2). The radiological evaluation was performed using a 3D FLAIR with sagittal reformats, coronal MPGR, STIR coronal 3-D T1 axial. The principal abnormalities were slight asymmetry in the hippocampal body with the left appearing smaller than the right and increased T2. No other abnormalities were identified. The radiologist concluded that the evidence was suggestive of left mesial temporal sclerosis. Additionally, multiple EEG monitoring sessions detected seizures originating from the left temporal lobe.

The individual feature maps for this patient are shown in Figure 2.5. IFMs corroborate the bilateral MRI and EEG clinical evaluations: The left-hemisphere map {L} indicates abnormal MD intensity in the left hippocampus. The corresponding report confirms that this abnormality is a hyperintensity (Appendix A.1). There are no hyper/hypointensity abnormalities in the right map {R}. The asymmetry map {KS} shows intensity asymmetry in the hippocampus (left>right) in both T2 and MD maps (multiple). There is also evidence of intensity asymmetry in the parahippocampal white matter on the T1 image (left>right). The volume asymmetry map {V} revealed left volume reduction in the hippocampus and right volume reduction in the superior-frontal cortex. This patient received a laterality score of -1 (perfect left).

2.3.2 Individual feature maps for a non-MTS patient

This patient was a 39 year-old, left-handed man (Patient #12 in Table 2.2) with a history of simple partial seizures during childhood. In his early 30s, he began experiencing stereotypical seizures with behavioural arrest, loss of awareness and buccal and bimanual automatisms. The MRI study was performed using diffusion and T2 gradient echo axial, FLAIR coronal, 3-D T1 axial pre and post gadolinium and T1 coronal pre and post gadolinium. Also abnormal T2 was evident in the right fusiform gyrus with some associated cortical thickening. A MRI diagnosis of dysembryoplastic neuroepithelial tumour (DNET) was favored, but the possibility of cortical dysplasia was also raised. Mild cerebellar atrophy as well as right frontal cortical atrophy were

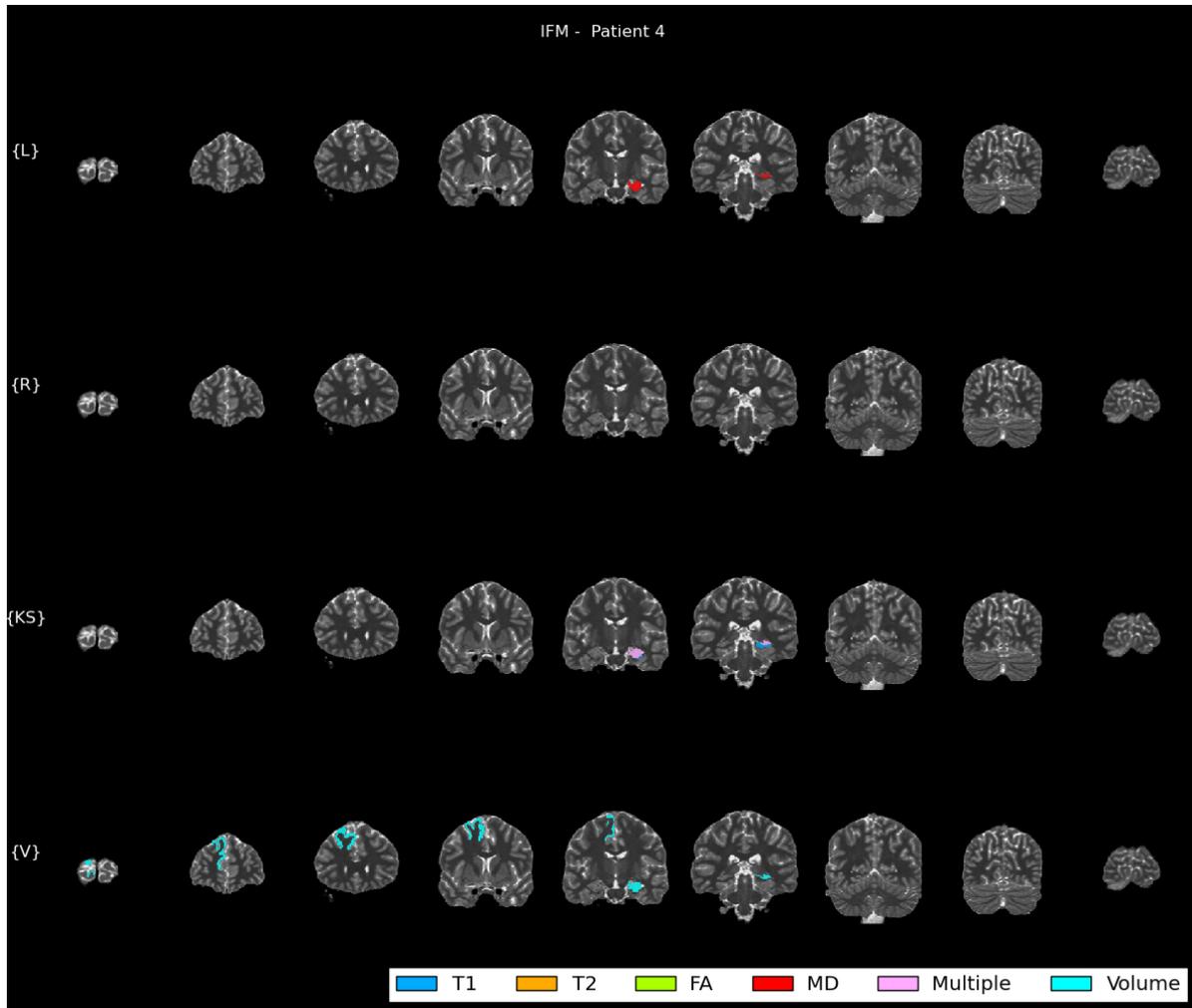


Figure 2.5: **IFM for a L-TLE patient with MTS.** The color code indicates the origin of the abnormality. There were no findings on the {R} map, therefore it is omitted here. This patient was categorized as a perfect left (laterality score=-1) because of the left asymmetry (left>right) in the parahippocampal region on T1, and the hippocampi in T2 and MD (multiple matches) shown on the {KS} map. (radiological convention is used, left is right)

also described. Scalp EEG recordings revealed independent bitemporal spikes and 2 seizures localized on the right mesial temporal region.

Figure 2.6 shows the IFMs for this patient, while the respective report is provided in Appendix A.2. Both the left {L} and right {R} individual feature maps showed focal abnormalities. The abnormalities in the left hemisphere were localized in the caudate (T2 hyperintensity) and the insular cortex (MD hyperintensity). Hyperintense T2 was found in the right frontal cortex, specifically in the pars triangularis (BA45 semantic tasks) and rostral middle frontal

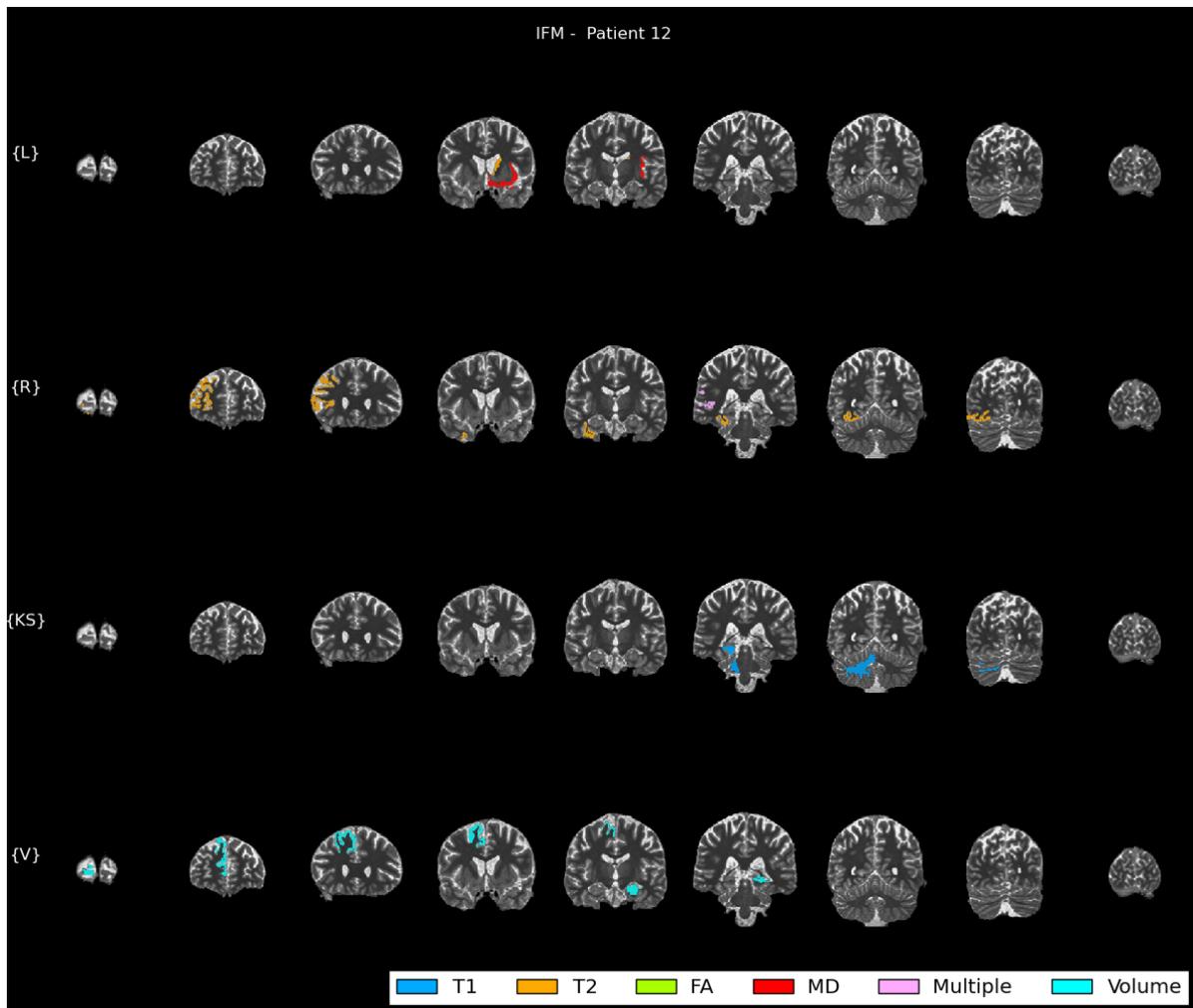


Figure 2.6: **IFMs for a R-TLE patient without MTS.** This patient received a laterality score of +1 (perfect right) due to the right asymmetry (right>left) in the parahippocampal region on the {KS} map. Radiological convention is used.

cortex (BA46 attention, short term memory). The banks of the superior temporal sulcus in the right temporal lobe appeared hyperintense on the MD and T2 images (multiple hits in the overlay). Notably, the fusiform cortex was hyperintense in MD, matching the presurgical radiological report which described an abnormal T2-weighted hyperintensity in the same region. The asymmetry map {KS} revealed intensity asymmetry (right>left) in the parahippocampal region as well as in the cerebellar white matter in T1. The volume map {V} showed changes in the superior-frontal cortex (left>right) and the hippocampi (right>left). The lateralization score based on temporal asymmetry for this patient was +1 (perfect right).

2.3.3 Group analysis

The group analysis was divided into two aspects: (i) differences between patients and controls in the frequency of intensity abnormalities (L, R, KS jointly); and (ii) differences regarding volume asymmetry. These analyses were performed using an independent samples t-test with a significance level α of 0.05. Differences between patients diagnosed with left TLE (L-TLE) and right TLE (R-TLE) were also studied using the same test. A second analysis was performed by grouping patients among those with no MTS (group of patients without mesial temporal sclerosis (MTS-)) and those with MTS (group of patients with mesial temporal sclerosis (MTS+)). Pairwise comparisons between controls, MTS- and MTS+ were carried out adjusting for multiple comparisons using the Bonferroni correction.

Ia - Intensity abnormalities across regions of interest

The frequency of intensity abnormalities was compared between patients and controls. For this test, all the extratemporal regions were grouped into extratemporal cortical and white matter regions. The amygdala, hippocampus, basal ganglia, temporal cortex, and temporal white matter regions were statistically different between the two groups (Figure 2.7a).

A similar comparison was performed, this time subdividing the patient group among MTS- and MTS+ patients (Figure 2.7b). Relevant differences appeared again in the amygdala, hippocampus, basal ganglia and temporal regions. However, this analysis revealed that intensity changes in the amygdala, hippocampus and basal ganglia are significantly more frequent in the MTS+ group (p -value <0.01). Abnormalities in the thalamus and ventricles were also identified (p -value <0.05). Interestingly, the frequency of intensity abnormalities was significantly higher in the MTS- group (p -value <0.01) with respect to controls, while it was less frequent in the MTS+ group. In contrast, the MTS+ group was also characterized by significantly higher occurrence of temporal white matter and extratemporal cortical abnormalities (p -value <0.05).

No statistical differences regarding the frequency of intensity abnormalities findings were found between L-TLE and R-TLE (Figure 2.7c).

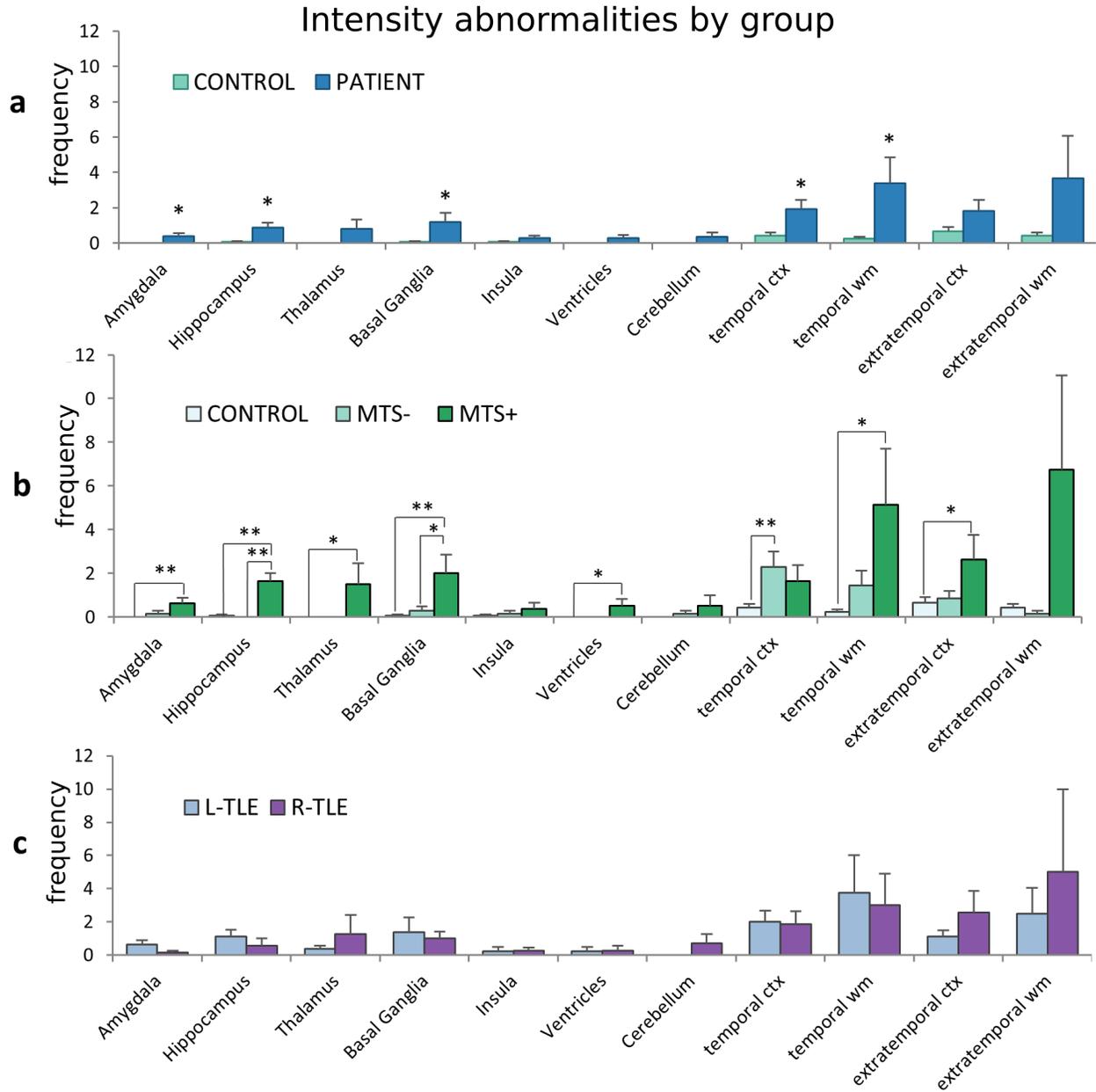


Figure 2.7: **Intensity abnormalities by group.** a) controls vs. patients; b) controls vs. MTS- vs. MTS+; c) L-TLE vs R-TLE. *($p < 0.05$), **($p < 0.01$)

Ib - Intensity abnormalities by image type

Intensity abnormalities were analyzed for each image source (T1, T2, FA, MD) and subject group (controls, MTS- and MTS+). A progression in asymmetry from controls with the lowest towards MTS+ with the highest was observed (Figure 2.8b). Notably, most of the asymmetry

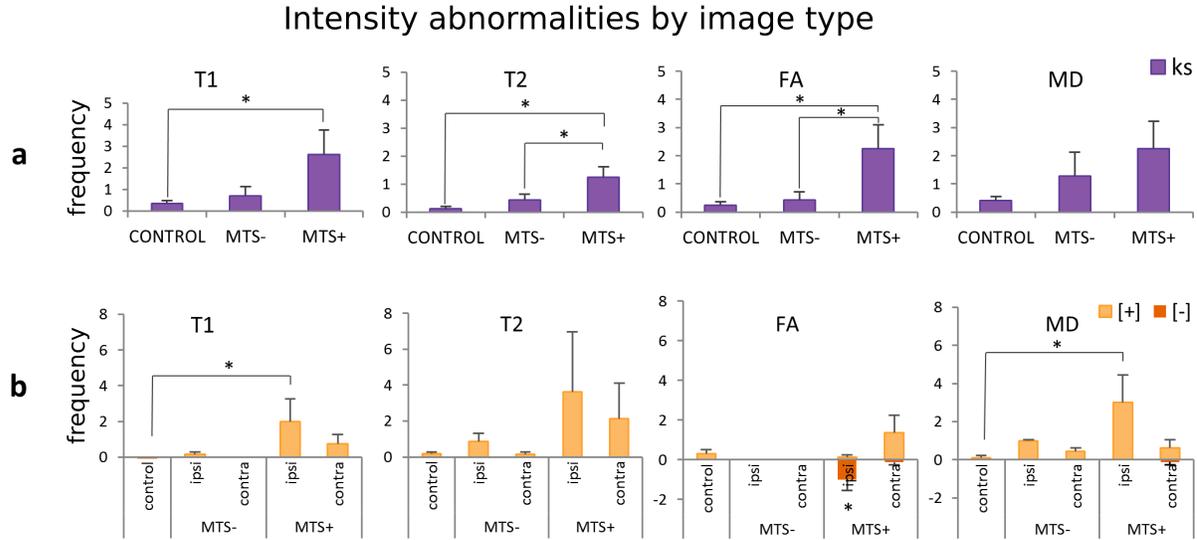


Figure 2.8: **Intensity abnormalities by image type.** a) *intensity asymmetry* (KS); b) *mean intensity* features organized as ipsilateral (ipsi) or contralateral (contra) instead of L/R. Hyperintensities are shown as [+] on the positive vertical axis and hypointensities as [-] on the negative vertical axis. *($p < 0.05$).

abnormalities in the MTS- group originated from MD. Significant differences (p -value <0.05) were found between controls and MTS+ on T1, T2 and FA; and between MTS- and MTS+ on T2 and FA. No significant differences were found between controls and MTS- patients.

An analogous analysis was performed comparing controls, MTS- and MTS+ with respect to the number of *mean intensity* features (L,R) which were reorganized as ipsilateral or contralateral to the hemisphere of the surgery. As shown in Figure 2.8b, the MTS+ group is characterized by ipsilateral hyperintensities on T1 and MD (p -value <0.05) and ipsilateral FA hypointensities (p -value <0.05). No statistical differences were found between MTS- patients and controls.

IIa - Volume asymmetry by region of interest

Patients were compared against controls in terms of *volume asymmetry*. Similar to previous tests, extratemporal regions were grouped. However, no comparison of volume asymmetry in white matter was performed. As shown in Figure 2.9a, only hippocampal volume asymmetry was statistically different between patients and controls ($p < 0.05$).

Subsequently, patients were divided between MTS- and MTS+ and compared against con-

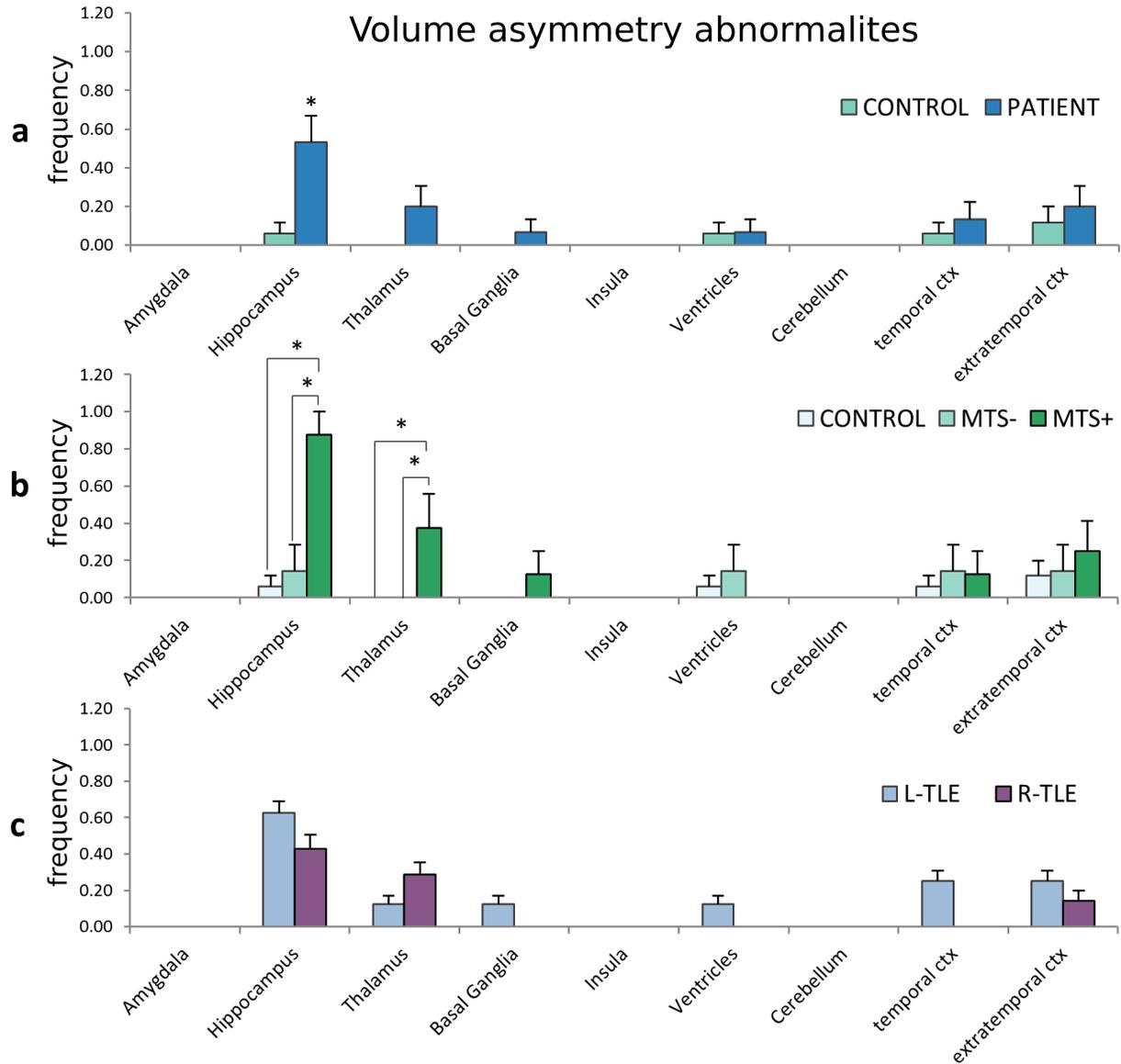


Figure 2.9: **Comparison of volume asymmetry abnormalities between groups** a) controls vs. patients; b) controls vs. MTS- vs. MTS+; c) L-TLE vs. R-TLE. *($p < 0.05$)

trols. Hippocampal volume asymmetry in MTS+ was significantly higher with respect to MTS- and control groups. A similar result was obtained in the thalamus (Figure 2.9b).

No statistically significant differences on volume asymmetry were found between L-TLE and R-TLE patients.

IIIb - Ipsilateral/contralateral volume reduction

Volume asymmetry results were reorganized as ipsilateral or contralateral volume reductions to the lesional side (L-TLE, R-TLE). Both L-TLE and R-TLE groups were characterized by a higher frequency of ipsilateral volume reductions. As shown in Figure 2.10a, ipsilateral volume reduction was statistically significant ($p < 0.05$) only in the L-TLE group.

A similar evaluation was performed comparing controls, MTS- and MTS+ groups. In this case, the MTS+ group was significantly different ($p < 0.01$) from the MTS- and control groups in frequency of ipsilateral volume reductions (Figure 2.10b).

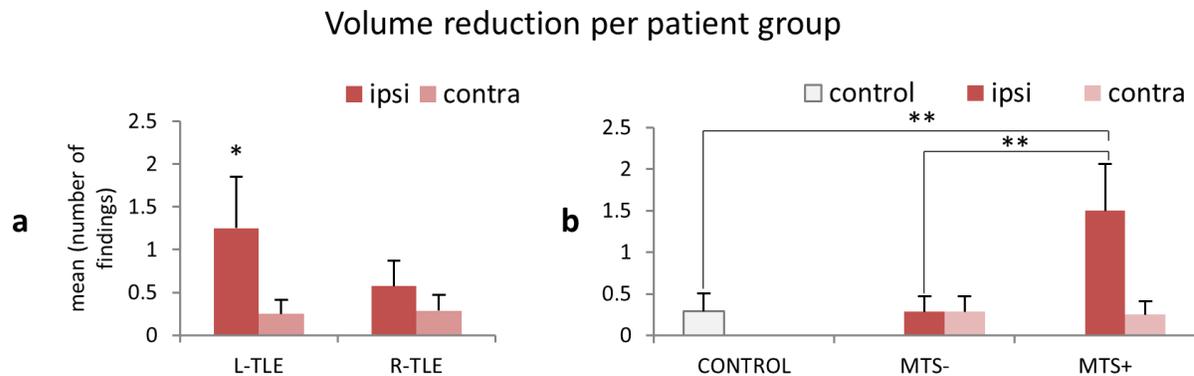


Figure 2.10: **Volume asymmetry (contralateral vs. ipsilateral)**. Frequency of *volume asymmetry* features organized into ipsilateral (ipsi) or contralateral (contra) to the lesional side; **a**: L-TLE vs R-TLE; **b**: controls vs. MTS- vs. MTS+. *($p < 0.05$), **($p < 0.01$)

Group-wise comparison of hippocampal volumes

Given that most of volume asymmetry differences were detected in the hippocampus, a direct group-wise comparison of volumes was performed, differentiating between ipsilateral and contralateral hippocampal volumes. Prior to the comparison, volumes were normalized using the total intracranial volume of each subject. A one-way ANOVA was performed comparing controls, MTS- and MTS+ showing a statistically significant difference in mean value of ipsilateral hippocampal volumes among the groups ($F = 16.96$, $p < 0.01$). Post-hoc pairwise comparisons showed a significant ipsilateral reduction of hippocampal volume for MTS+ with respect to the MTS- and control groups ($p < 0.01$). No significant differences were found between controls

and MTS- groups on ipsilateral volume reduction. No significant differences were obtained contralaterally.

2.4 Discussion

2.4.1 Study design considerations and limitations of this work

Individual Feature Maps (IFMs) allow comparisons between a patient and a set of healthy subjects to determine regions of interest that are significantly different in the patient with respect to the control group. The nature of the underlying statistical test is to identify *patient-specific abnormalities* rather than results that could be extrapolated to the entire patient population using statistics, as achieved by techniques such as voxel-based morphometry [28].

Another important consideration is that IFMs are sensitive to any departure from the control group distribution and not only to those that can be correlated to TLE. For example, differences in ventricular intensity were found for the MTS+ group (Figure 2.7b), however, one of the patients accounting for such differences was diagnosed with ventricular enlargement and periventricular leukomalacia (#8 on Table 2.2). Therefore, interpretation of IFMs should consider the clinical history and other findings accounting for possible *secondary pathologies* that may appear in addition to TLE. Also, it has been shown that MRI abnormalities associated with TLE can be common in healthy individuals [29]. Therefore, any finding should be examined in clinical context. In particular, no assumptions regarding the level of asymmetry in the control cohort are made. Instead, a baseline for normal asymmetry is established from the control group and it is then applied to evaluate the rarity of individual patient measurements.

I consider that the selection of the control group is an important factor for the applicability of IFMs in similar studies. The control group should match the type of patient (for example, pediatric patients must be compared against similar-aged control group). A large variance of the control group must be avoided to minimize false negatives.

Although a larger patient cohort is not required for individual patient analysis, it could improve the applicability of *a posteriori* group-wise analysis of the kind presented in this chapter.

Though the number of extratemporal (cortical and white matter) intensity abnormalities was higher for R-TLE with respect to L-TLE (Figure 2.7c), no statistical significance was attained. Similarly, a more detailed analysis looking for differences and diagnostic clues for the elusive MTS- group would be possible with a larger MTS- patient sample.

2.4.2 On the role of single-subject statistics

Single-subject t-tests are fundamental to the acquisition of IFMs. There are two considerations (or limitations) regarding the use of these tests in small samples: a) departures from normality and b) statistical power.

Crawford et al.[22] analyzed the effects of lack of normality on normative samples (control group) on single-subject t-tests and conclude that this test behaves better than a regular z-score test (more commonly used by clinicians) for small datasets ($N < 20$) and that departures from normality have low to moderate effects in the test score. If there is concern over the skewness and/or kurtosis of the normative sample, then a smaller significance level than the commonly used 0.05 (i.e. 0.02) could give a higher degree of confidence that the result is not a non-normality artifact, effectively reducing Type I error inflation. A practical consideration is that while patient images are difficult to obtain (patient availability and eligibility), it is much easier to image control volunteers. Therefore, an increased volunteer cohort should ameliorate the quality of the probability distribution.

The second consideration is the statistical power on small datasets. In theory, to increase power the significance level of the test could be less severe ⁶. This would effectively reduce the number of false negative (Type II errors), increasing statistical power with the trade-off of increasing the number of false positives (Type I errors). In practice, this does not improve reliability. In our study of TLE, as in most statistical neuroimaging studies, the *true effect size is commonly unknown* [30]. In other words, it is unknown how large a difference between patients and controls needs to be, in order to be declared a pathology. Nonetheless, it is important to note that unlike two-sample t-tests where inferences on a patient group are sought, the goal

⁶by changing the significance threshold (larger p-value)

of a single-subject t-test is to obtain an estimate of the rarity of a feature *for a given subject* with respect to a normative distribution (control group). In other words, the null hypothesis of single-subject t-tests is that the subject being tested belongs to the normative population (control group). Thus, statistically significant results in this scenario, represent regions where the subject departs from the control distribution beyond a given statistical threshold set for the test, and no inferences are made over the patient population.

2.4.3 Comparison with similar studies

The analysis of T1 in the context of TLE has only recently been explored [31]. Our results showed T1 hyperintensities were significantly more common in MTS+ patients ($p < 0.05$), which could be an indication of gliosis and subsequent remyelination (higher signal due to myelin), or iron accumulation associated with microgliosis [32]. More studies are required to investigate the effectiveness of T1 measurements in the study of TLE as well as its correlation with histological findings.

Similarly to ROI analyses performed by Pell et al. [25] and VBR analyses by Mueller et al. [26], ipsilateral hippocampal T2 hyperintensities in MTS+ patients were found (Figure 2.8b). However, since our study included several image types, it was assessed that MD asymmetry was comparatively more frequent. In addition, MD hyperintensities were found in the ipsilateral temporal lobe of MTS+ patients. This type of finding has also been reported by Shon et al. [11].

Similar to our results, a pattern of ipsilateral FA hypointensity and MD hyperintensity in MTS+ patients (Figure 2.8b) was evidenced by Focke et al. [10] on 33 MTS+ patients. Significant FA reductions in patients with respect to controls were also found, similarly to Ahmadi et al. [9]. In parallel to Liu et al. [33], temporal white matter changes were significantly more frequent in MTS+ than in MTS-. In agreement with Gross et al. [34], no correlation between FA changes and surgical outcome were found.

It has been suggested that DTI abnormalities are not as extensive or as severe in MTS- patients as they are in MTS+ patients [35]. This tendency was observed in our results, with more

frequent intensity abnormalities in the MTS+ with respect to the MTS-. However, statistical significance was not achieved in our cohort (Figure 2.8a and b, FA and MD plots). In general, our results show that volume and intensity abnormalities are present to a lesser degree in MTS- patients. Nonetheless, intensity changes in the temporal cortex were significantly more frequent ($p < 0.01$) in this group (Figure 2.7b). The subsequent analysis, revealed that most of these abnormalities originate from MD images (Figure 2.8a and b). This suggests that the analysis of the temporal cortex using MD images could be key for the study of non-lesional cases. In fact Keller et al. [12] have studied this scenario in 10 non-lesional TLE patients, obtaining patterns of FA reduction and MD increase. However, they conclude that larger cohorts are required to obtain more reliable results.

Though extratemporal changes were slightly more common in our R-TLE group, no statistical significance was obtained.

The group-wise analysis of volume differences was in agreement with the results shown by Seidenberg et al. [27] and Cohan et al. [8]: hippocampal volume changes were statistically significant between patients and controls. In addition, satisfactory results were obtained using hippocampal volume to distinguish between controls, MTS- and MTS+. Also, extratemporal volume asymmetry on the cortex was found in TLE patients, though it was not significantly higher than that found in our control cohort (Figure 2.9a). Also, in agreement with Pell et al. [1], volume differences between patients and controls were seen in the hippocampus and thalamus (Figure 2.9b). Moreover, our study seems to indicate that both intensity and volumetric thalamic changes are significantly more frequent in MTS+ patients, which suggests that the thalamus is an interesting region to analyze as a hub in the pathologic network of TLE [36, 37]

2.4.4 Clinical Relevance

Our multi-feature evaluation approach identified hippocampal involvement in all the patients with MTS (Table 2.2). Patient #6 who had been identified as a possible MTS in the presurgical MRI was later on confirmed negative for MTS by histopathology, which only revealed gliosis. Consistently with the pathological validation, the IFMs did not report any intensity

or or volumetric abnormalities for this patient. No hippocampal involvement was reported in the MTS- with the exception of patient #12, where the volume IFM {V} reported volume reduction in the left hippocampus. Also, in the MTS- group, significant intensity changes were detected in the temporal cortex with respect to the control group (p -value < 0.01) as shown in Figure 2.7b. Focusing on the detection of quantitative intensity changes on the temporal cortex could contribute to a better understanding of non-lesional cases.

Using the *laterality score*, which measures intensity asymmetry in the temporal lobes, 13 out of 15 patients were lateralized correctly, with the remaining three receiving a lateralization score of zero. These subjects had normal presurgical MRIs (Table 2.2, #13-15) and our method did not find sufficient intensity asymmetry in their respective T1, T2, MD, and FA images.

No clinical correlation was found between the features evaluated in this study and the Engel classification for surgical outcome.

2.5 Conclusion

This chapter introduced a novel method that demonstrates the integration of high-resolution relaxometry (1mm) and DTI to perform TLE *patient-specific* analyses. The IFMs integrate heterogeneous data signaling regions that may be of clinical interest based on their rarity, and present these results in a simple and efficient manner to clinicians.

Our method detected hippocampal involvement in all of the MTS cases. More importantly, abnormal-looking regions were clearly identified in non-lesional and MTS- cases, pointing to changes in the temporal cortex as a key to improve the understanding of these cases. Also, the IFMs provided the correct lateralization of all patients (L-TLE or R-TLE) with the exception of three patients for whom a laterality score of zero coincided with normal presurgical MRIs (Table 2.2).

The subsequent group-level analysis revealed expected patterns such as ipsilateral hyperintensities in T1, T2 and MD accompanied by ipsilateral hypointensities in FA in the MTS+ group. The description of such patterns would have not been possible without the integrative

approach of this work encompassing quantitative relaxometry and diffusion tensor imaging. Individual epilepsy diagnosis based on MR images can be improved by the inclusion of these type of sequences in current imaging protocols. Also, IFMs along with the lateralization score have the potential to supplement invasive sub-dural or deep electrode monitoring for improved localization. The reliability of the presented results can be improved by increasing the size of the control group (representing the healthy population), as well as by a larger patient cohort for further individual validation and group-wise analyses.

Bibliography

- [1] G. S. Pell, R. S. Briellmann, H. Pardoe, D. F. Abbott, G. D. Jackson, Composite voxel-based analysis of volume and T2 relaxometry in temporal lobe epilepsy, *Neuroimage* 39 (3) (2008) 1151–1161.
- [2] S. S. Keller, N. Roberts, Voxel-based morphometry of temporal lobe epilepsy: An introduction and review of the literature, *Epilepsia* 49 (5) (2008) 741–757.
- [3] N. Kemmotsu, H. M. Girard, B. C. Bernhardt, L. Bonilha, J. J. Lin, E. S. Tecoma, V. J. Iragui, D. J. Hagler, E. Halgren, C. R. McDonald, MRI analysis in temporal lobe epilepsy: cortical thinning and white matter disruptions are related to side of seizure onset, *Epilepsia* 52 (12) (2011) 2257–2266.
- [4] N. Memarian, P. M. Thompson, J. Engel, R. J. Staba, Quantitative analysis of structural neuroimaging of mesial temporal lobe epilepsy, *Imaging in medicine* 5 (3) (2013) 219–235.
- [5] N. Bernasconi, D. Kinay, F. Andermann, S. Antel, A. Bernasconi, Analysis of shape and positioning of the hippocampal formation: an mri study in patients with partial epilepsy and healthy controls, *Brain* 128 (10) (2005) 2442–2452.
- [6] S. S. Keller, J.-C. Schoene-Bake, J. S. Gerdes, B. Weber, M. Deppe, Concomitant fractional anisotropy and volumetric abnormalities in temporal lobe epilepsy: Cross-sectional evidence for progressive neurologic injury, *PloS one* 7 (10) (2012) e46791.
- [7] M. Esmailzadeh, H. Soltanian-Zadeh, K. Jafari-Khouzani, Mesial temporal lobe epilepsy lateralization using SPHARM-based features of hippocampus and SVM, in: *Proc. SPIE*, Vol. 8314, 2012, pp. 83144H–83144H–10.
- [8] A. Coan, B. Kubota, F. Bergo, B. Campos, F. Cendes, 3T MRI Quantification of Hippocampal Volume and Signal in Mesial Temporal Lobe Epilepsy Improves Detection of Hippocampal Sclerosis, *American Journal of Neuroradiology*.

- [9] M. Ahmadi, D. Hagler, C. McDonald, E. Tecoma, V. Iragui, A. Dale, E. Halgren, Side matters: Diffusion tensor imaging tractography in left and right temporal lobe epilepsy, *American Journal of Neuroradiology* 30 (9) (2009) 1740–1747.
- [10] N. K. Focke, M. Yogarajah, S. B. Bonelli, P. A. Bartlett, M. R. Symms, J. S. Duncan, Voxel-based diffusion tensor imaging in patients with mesial temporal lobe epilepsy and hippocampal sclerosis, *Neuroimage* 40 (2) (2008) 728–737.
- [11] Y.-M. Shon, Y.-I. Kim, B.-B. Koo, J.-M. Lee, H. J. Kim, W. J. Kim, K. J. Ahn, D. W. Yang, Group-specific regional white matter abnormality revealed in diffusion tensor imaging of medial temporal lobe epilepsy without hippocampal sclerosis, *Epilepsia* 51 (4) (2010) 529–535.
- [12] S. S. Keller, T. Ahrens, S. Mohammadi, J. S. Gerdes, G. Möddel, C. Kellinghaus, H. Kugel, B. Weber, E. B. Ringelstein, M. Deppe, Voxel-based statistical analysis of fractional anisotropy and mean diffusivity in patients with unilateral temporal lobe epilepsy of unknown cause, *Journal of Neuroimaging*.
- [13] S. C. Deoni, B. K. Rutt, T. M. Peters, Rapid combined T1 and T2 mapping using gradient recalled acquisition in the steady state, *Magnetic Resonance in Medicine* 49 (3) (2003) 515–526.
- [14] S. C. L. Deoni, T. M. Peters, B. K. Rutt, High-resolution T1 and T2 mapping of the brain in a clinically acceptable time with DESPOT1 and DESPOT2, *Magnetic Resonance in Medicine* 53 (1) (2005) 237–241.
- [15] S. C. Deoni, High-resolution T1 mapping of the brain at 3T with driven equilibrium single pulse observation of T1 with high-speed incorporation of rf field inhomogeneities (DESPOT1-HIFI), *Journal of Magnetic Resonance Imaging* 26 (4) (2007) 1106–1111.
- [16] S. C. Deoni, Transverse relaxation time (T2) mapping in the brain with off-resonance correction using phase-cycled steady-state free precession imaging, *Journal of Magnetic Resonance Imaging* 30 (2) (2009) 411–417.

- [17] M. Jenkinson, S. Smith, A global optimisation method for robust affine registration of brain images, *Medical image analysis* 5 (2) (2001) 143–156.
- [18] M. Beg, M. I. Miller, A. Troune, L. Younes, Computing large deformation metric mappings via geodesic flows of diffeomorphisms, *International Journal of Computer Vision* 61 (2) (2005) 139–157.
- [19] H. Huang, C. Ceritoglu, X. Li, A. Qiu, M. I. Miller, P. C. van Zijl, S. Mori, Correction of B0 susceptibility induced distortion in diffusion-weighted images using large-deformation diffeomorphic metric mapping, *Magnetic Resonance Imaging* 26 (9) (2008) 1294 – 1302.
- [20] S. C. Deoni, B. K. Rutt, T. M. Peters, Synthetic T1-weighted brain image generation with incorporated coil intensity correction using DESPOT1, *Magnetic Resonance Imaging* 24 (9) (2006) 1241 – 1248.
- [21] J. Crawford, D. C. Howell, Comparing an individual’s test score against norms derived from small samples, *The Clinical Neuropsychologist* 12 (4) (1998) 482–486.
- [22] J. R. Crawford, P. H. Garthwaite, A. Azzalini, D. C. Howell, K. R. Laws, Testing for a deficit in single-case studies: Effects of departures from normality, *Neuropsychologia* 44 (4) (2006) 666 – 677.
- [23] J. R. Crawford, P. H. Garthwaite, D. C. Howell, On comparing a single case with a control sample: An alternative perspective, *Neuropsychologia* 47 (13) (2009) 2690 – 2695.
- [24] J. R. Crawford, P. H. Garthwaite, Single-case research in neuropsychology: A comparison of five forms of t-test for comparing a case to controls, *Cortex* 48 (8) (2012) 1009–1016.
- [25] G. S. Pell, R. S. Briellmann, A. B. Waites, D. F. Abbott, G. D. Jackson, Voxel-based relaxometry: a new approach for analysis of T2 relaxometry changes in epilepsy, *Neuroimage* 21 (2) (2004) 707–713.

- [26] S. G. Mueller, K. D. Laxer, N. Schuff, M. W. Weiner, Voxel-based T2 relaxation rate measurements in temporal lobe epilepsy (tle) with and without mesial temporal sclerosis, *Epilepsia* 48 (2) (2007) 220–228.
- [27] M. Seidenberg, K. G. Kelly, J. Parrish, E. Geary, C. Dow, P. Rutecki, B. Hermann, Ipsilateral and contralateral MRI volumetric abnormalities in chronic unilateral temporal lobe epilepsy and their clinical correlates, *Epilepsia* 46 (3) (2005) 420–430.
- [28] J. Ashburner, K. J. Friston, Voxel-based morphometry - the methods, *Neuroimage* 11 (6) (2000) 805–821.
- [29] A. Labate, A. Gambardella, U. Aguglia, F. Condino, P. Ventura, P. Lanza, A. Quattrone, Temporal lobe abnormalities on brain MRI in healthy volunteers a prospective case-control study, *Neurology* 74 (7) (2010) 553–557.
- [30] K. S. Button, J. P. Ioannidis, C. Mokrysz, B. A. Nosek, J. Flint, E. S. Robinson, M. R. Munafò, Power failure: why small sample size undermines the reliability of neuroscience, *Nature Reviews Neuroscience* 14 (5) (2013) 365–376.
- [31] A. R. Khan, M. Goubran, S. de Ribaupierre, R. R. Hammond, J. G. Burneo, A. G. Parent, T. M. Peters, Quantitative relaxometry and diffusion MRI for lateralization in mts and non-mts temporal lobe epilepsy, *Epilepsy research* 108 (3) (2014) 506–516.
- [32] S. Shoham, M. Youdim, Iron involvement in neural damage and microgliosis in models of neurodegenerative diseases., *Cellular and molecular biology (Noisy-le-Grand, France)* 46 (4) (2000) 743–760.
- [33] M. Liu, L. Concha, C. Lebel, C. Beaulieu, D. W. Gross, Mesial temporal sclerosis is linked with more widespread white matter changes in temporal lobe epilepsy, *NeuroImage: clinical* 1 (1) (2012) 99–105.
- [34] D. W. Gross, L. Concha, C. Beaulieu, Extratemporal white matter abnormalities in mesial

temporal lobe epilepsy demonstrated with diffusion tensor imaging, *Epilepsia* 47 (8) (2006) 1360–1363.

[35] D. W. Gross, Diffusion tensor imaging in temporal lobe epilepsy, *Epilepsia* 52 (2011) 32–34.

[36] B. C. Bernhardt, N. Bernasconi, H. Kim, A. Bernasconi, Mapping thalamocortical network pathology in temporal lobe epilepsy, *Neurology* 78 (2) (2012) 129–136.

[37] S. S. Keller, J. O’Muircheartaigh, C. Traynor, K. Towgood, G. J. Barker, M. P. Richardson, Thalamotemporal impairment in temporal lobe epilepsy: A combined MRI analysis of structure, integrity, and connectivity, *Epilepsia* 55 (2) (2014) 306–315.

Chapter 3

TLE detection using support vector machines in multi-parametric quantitative MR imaging

3.1 Introduction

The detection of MRI abnormalities that can be associated with seizures in the study of temporal lobe epilepsy (TLE) is a challenging task. In many cases, patients with a record of epileptic activity do not present any discernible MRI findings. Recent studies have examined the possibility of improving TLE detection using SVMs on MRI data. For example Focke et al. [1] show correct patient lateralization (left vs. right seizure onset) using SVMs on T1-weighted and DTI data. In addition, Keihaninejad et al. [2], demonstrate the identification of TLE cases with hippocampal atrophy from cases without it using SVM on regional volumes obtained from T1-weighted MRI.

This chapter explores the integration of quantitative relaxometry and diffusion tensor imaging with support vector machines (SVM) aiming to improve TLE detection. Quantitative relaxometry produces images that are independent from acquisition parameters, thus having the potential for better describing tissue abnormalities. Similarly, as mentioned in chapter 1, dif-

fusion tensor imaging has identified patterns of disruption in the white matter in TLE patients. In this chapter machine learning models on these promising type of images are developed and compared to determine an optimal model for classification.

3.2 Methods

3.2.1 Overview

All subjects in this study underwent an imaging protocol approved by the Office of Research Ethics at Western University (Canada) ¹. The imaging protocol comprised DESPOT1-HIFI, DESPOT2-FM [3, 4] and DTI sequences, resulting image data being processed to obtain four different quantitative maps: T1, T2, fractional anisotropy (FA) and mean diffusivity (MD). Anatomical atlas-based labeling was used to define ROIs on each one of the quantitative maps and subsequently to measure and extract regional features.

Given that the number of features exceeded the number of subjects (936 features, 36 subjects), two different feature selection methods were explored to discard irrelevant or uninformative features. Then, the feature space was further reduced using principal component analysis (PCA).

An SVM was trained/tested on the filtered feature space using a leave-one-out cross-validation strategy (LOOCV), where the SVM was tasked with predicting the label (patient or control) for the omitted subject (Figure 3.1).

Once all subjects were evaluated, sensitivity, specificity and classification accuracy were measured and reported. This procedure was repeated for each *classification model* obtained by the combination of the following elements:

- the image that originates the features (T1, T2, FA, MD, or all combined)
- the method to select features
- the cardinality of the requested feature set [K]

¹The ethics approval document is available in Appendix B

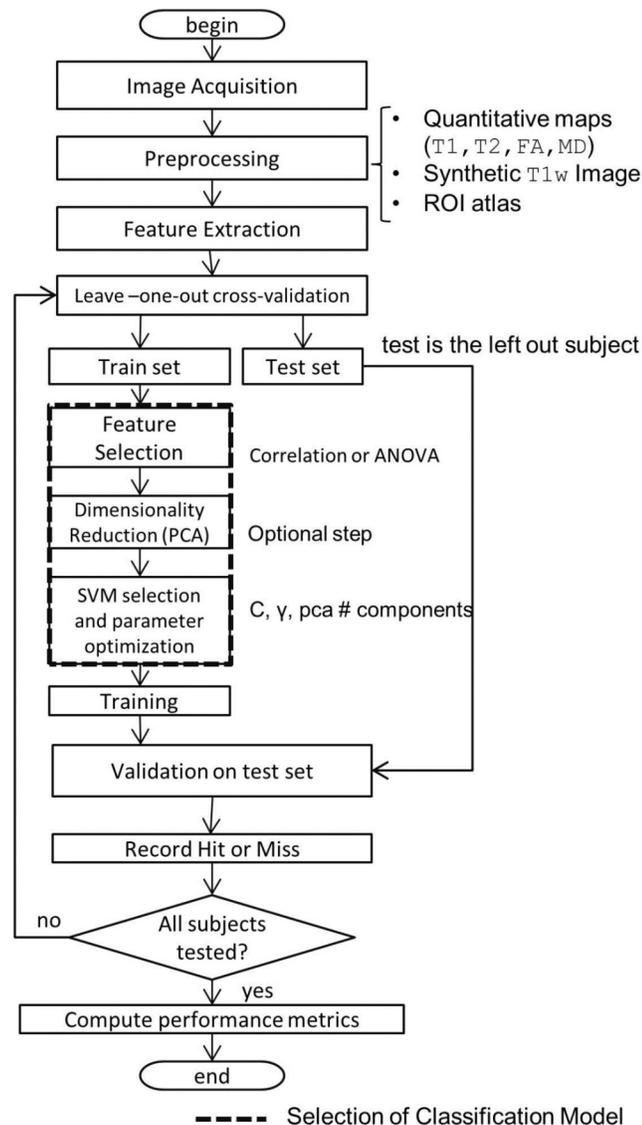


Figure 3.1: **Method overview.** The preprocessing and feature extraction steps are the same for all the experiments and therefore they are executed only once. Each of the experiments described in this chapter corresponds to different instances of the leave-one-out cross-validation (LOOCV) loop. The model selection stage determines the feature space, its dimension (PCA-reduced or not), and the parameters that constitute the classification model. 8 models in total are evaluated for each classification experiment. These models are evaluated using performance metrics that are collected at the end of each LOOCV loop.

- the use of PCA to reduce the feature space
- the type of SVM

A detailed comparison among classification models is reported in the results section along the best classification scenarios. Specific recommendations regarding the optimal model are given in the discussion section. In addition, an analysis of the elements constituting a classification model and their influence in the classifiers performance is also discussed. Finally, the features relevant for classification are analyzed and their clinical significance is considered.

3.2.2 Participants

Thirty-six individuals participated in this study, 19 of whom were control volunteers (age 32 ± 10 , 12 male, 7 female) and 17 TLE patients (age 35 ± 10 , 8 male, 9 female). All the patients had lateralizable seizures (confirmed by EEG) and all of them were eligible for temporal lobectomy (9 left, 8 right). Preoperative MRI and post-surgical pathology confirmed the presence of Mesial Temporal Sclerosis (MTS) in 8 of them.

3.2.3 Imaging Protocol and Preprocessing

Subjects were scanned (presurgically in the case of patients) using a 3T MR scanner (GE Discovery MR750) with whole brain, 1mm isotropic DESPOT1-HIFI and DESPOT2-FM, T1 and T2 mapping sequences [3, 4] respectively, optimized for imaging at 3T [5]. Two SPGR images were acquired (flip angles of 4° and 18°) along with an inversion-prepared spoiled gradient recalled acquisition in the steady state (SPGR) to calculate a quantitative T1 map. Five balanced steady state free-precession (bSSFP) images were acquired with phase cycling (flip angles of $15^\circ, 35^\circ, 60^\circ$) to estimate a quantitative T2 map using the DESPOT2-FM procedure [6]. All images were co-registered to the 18° flip angle SPGR image using the FLIRT registration tool [7] from the FSL software (<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>) to account for motion between scans prior to the computation of the T1 and T2 maps.

A DTI sequence was also acquired with the following parameters: 2.4mm isotropic, 41 directions, b-value =1000, 4 non-weighted (b=0) volumes. Non-linear distortions were corrected by deformable registration of the average of the b=0 volumes to the undistorted T1 map, using a diffeomorphic registration method [8, 9]. Eddy current correction and diffusion tensor estimation was performed using FSL's diffusion toolbox FDT. Maps of fractional anisotropy (FA) and mean diffusivity (MD) were transformed and re-sampled to the coordinate system defined by the 1mm isotropic T1 map. Synthetic T1-weighted images, with inherent bias-field correction, were generated from the T1 maps [10] and used in place of directly acquired T1-weighted images for subsequent segmentation.

This preprocessing stage yielded four co-registered quantitative maps (T1, T2, FA and MD) for each subject.

3.2.4 Feature Extraction

Volumetric segmentation of the synthetic T1-weighted images was performed with Freesurfer (<http://surfer.nmr.mgh.harvard.edu/>). This processing included removal of non-brain tissue using a hybrid watershed/surface deformation procedure, automated Talairach transformation, segmentation of the subcortical white matter and deep gray matter volumetric structures (including hippocampus, amygdala, caudate, putamen, ventricles). Once the cortical models were completed, the cerebral cortex was parcellated into regions based on gyral and sulcal structures [11]. This segmentation produced bilateral regions of interest for every subject in subject image space. Total of 35 cortical, 34 white matter and 9 subcortical regions (78 in all) were identified per hemisphere. The segmented ROIs were used to extract features from the quantitative maps (T1, T2, FA, and MD) for each subject (Figure 3.2).

Two types of features were extracted: The ROI *mean intensity*, and the *asymmetry* between correspondent left and right ROIs (i.e., left and right hippocampus). The asymmetry was expressed as the non-parametric two-sample Kolmogorov-Smirnov statistical score between left-right region pairs. For each subject, a vector of 936 features was created, accounting for 624 mean intensity features (156 ROIs, 4 image sources) and 312 intensity difference features (78

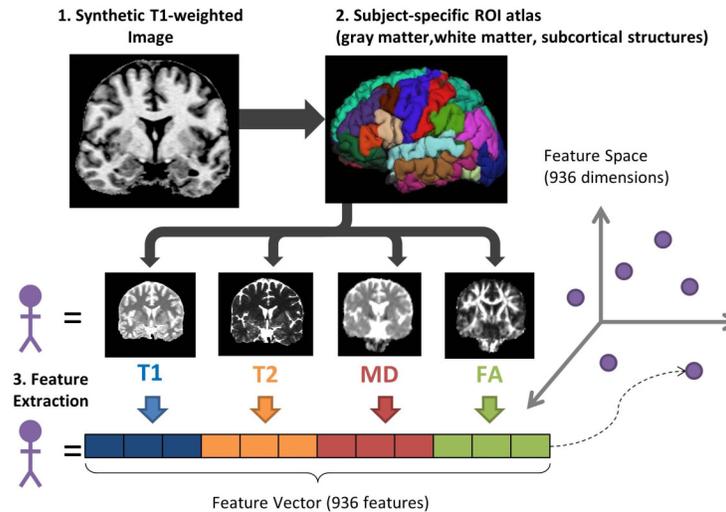


Figure 3.2: **Image preprocessing and feature extraction steps.** A subject-specific ROI atlas is employed to extract features from the four different quantitative maps as shown. The resulting feature vector represents each subject in a highly-dimensional feature space. Upon this space is that feature selection algorithms are used to discriminate features that are relevant for classification.

left-right region pairs, 4 image sources). Each feature vector was labelled according to the respective subject ground truth (i.e. patient or control). Prior to using these feature vectors as input for the SVM, irrelevant features were discarded.

3.2.5 Support Vector Machines

Linear SVMs are used in scenarios where the features that discriminate between subject groups are linearly separable. Linear SVMs estimate the hyperplane or *decision boundary* that provides the largest margin (separation) between the two groups [12]. In contrast, non-linear SVMs such as *radial basis function* SVMs provide a non-linear boundary using the *kernel trick* as described in [13] which effectively transforms the non-linear space into a space that has a higher number of dimensions but where the features are linearly separable. The *scikit-learn* library [14] was used to implement the SVM classifiers in this work. Both linear and non-linear support vector machines were used for patient detection (Figure 3.3).

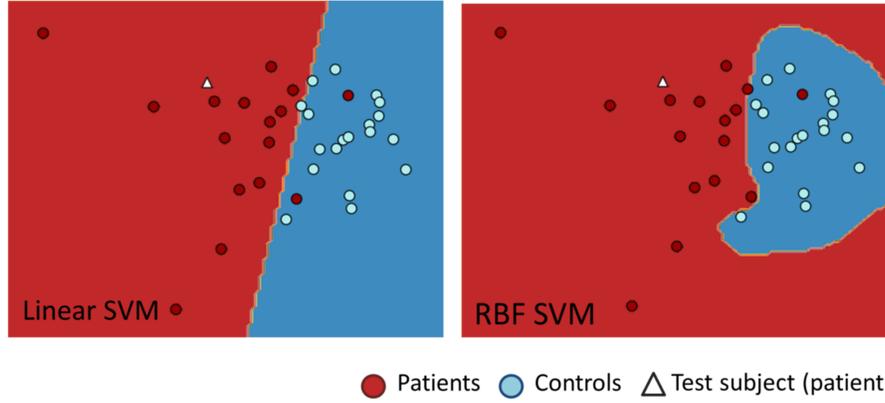


Figure 3.3: **Linear vs. non-linear SVMs.** This example shows the decision boundaries of a Linear SVM and a radial basis function(RBF) SVM for the same feature space. The scenario shown is the automatic classification between patients and controls (Experiment I). Here, one of the patients has been selected as the testing set. The classifier is trained using all the other subjects (training set). The patient used as the test set appears to be correctly classified by both the Linear and the RBF SVMs.

3.2.6 Experiments

Each subject was assigned one of two possible labels: *patient* or *control*. The SVM was trained using feature vectors to distinguish between these two groups. A leave-one-out cross-validation approach (LOOCV) was used for training/testing. One subject was left out from the training stage while the remainder were used for training (i.e. to compute the decision boundary). Then, the classifier was asked to predict the correct label for the excluded subject. This procedure iterated until all subjects had been classified. Afterwards, the *sensitivity*, *specificity* and *accuracy* were calculated as follows:

$$sensitivity = \frac{tp}{tp + fp} \quad (3.1)$$

$$specificity = \frac{tc}{fp + tc} \quad (3.2)$$

$$accuracy = \frac{tp + tc}{p + c} \quad (3.3)$$

where tp = correctly identified patients, tc = correctly identified controls, fp = controls mis-

classified as patients, fc = patients misclassified as controls, p = total number of patients, c = total number of controls.

Patients were subsequently split into those with left temporal lobe seizures and those with right temporal lobe seizures (L-TLE and R-TLE respectively). The same training/testing procedure was repeated to assess the detectability of each patient subgroup.

3.2.7 Classification Models

The outcome of an SVM is influenced by the feature space over which it operates (extrinsic element) as well as by the type of SVM (intrinsic element). In this work, we define a *classification model* as the combination of *extrinsic* and *intrinsic* elements selected for the operation of the SVM.

Each model is built in three steps as shown in Figure 3.1: One of two possible feature selection methods (correlation-based or ANOVA-based); the option or not of performing dimensionality reduction with PCA; and the type of SVM (linear or non-linear). Eight different classification models were evaluated:

- *correlation-svm-linear*: Features are selected with a correlation-based method and a linear SVM runs on this space.
- *correlation-pca-svm-linear*: After the correlation-based feature selection, the dimension of the resulting feature space is reduced using PCA. A linear SVM is used for classification.
- *correlation-svm-rbf*: Features are selected with a correlation-based method and the SVM used is non-linear (radial basis functions).
- *correlation-pca-svm-rbf*: The dimension of the feature space obtained with correlation-based selection is simplified using PCA. The employed SVM is non-linear.
- *anova-svm-linear*: Features are selected using an ANOVA-based approach. The SVM used is linear.

- *anova-pca-svm-linear*: Features are selected using an ANOVA-based approach. The resulting feature space is transformed using PCA. The SVM used is linear.
- *anova-svm-rbf*: Features are selected using an ANOVA-based approach. The SVM is non-linear.
- *anova-pca-svm-rbf*: Similar to the previous case, but the resulting feature space is reduced with PCA.

The following sections describe the steps taken to build a classification model, while the results section compares models based on their performance. Finally, the best and worst models considering feature selection, PCA, cardinality and computational performance elements are discussed and an analysis of relevant, stable features and their clinical significance is presented.

3.2.8 Feature Selection

Feature selection algorithms identify features that are pivotal for classification [15], and optimize the accuracy of machine learning algorithms (such as SVMs), while avoiding overfitting of the data [16].

In theory, feature selection should be performed on data that are independent of both the training and testing sets [17, p. 222]. In practice, this is not always feasible, particularly when the number of features is much larger than the number of subjects, as is the case in our work. To address this issue, our feature selection step subsamples the training set, creating leave-one-subject-out folds and then scores features on each resulting fold. After this, a voting algorithm examines the agreement among the folds, and decides which features belong to the final feature set. This approach reuses the data efficiently and reduces the outlier effect (performance reduction attributable to a subject).

Training Set Subsampling

The training set was subsampled by leaving one subject out every time, creating M folds for a training set of size M (Figure 3.4). On each fold, one of two possible feature selection methods

(correlation or ANOVA-based) was evaluated. Each method began by assigning a score to each of the 936 features for each subject in the current fold. Then, using the *cardinality parameter* K , set by the user, the respective method selected the top $2K$ features based on their scores. Finally, after all folds were evaluated, the voting algorithm decided on the final K features by reviewing the agreement among the folds.

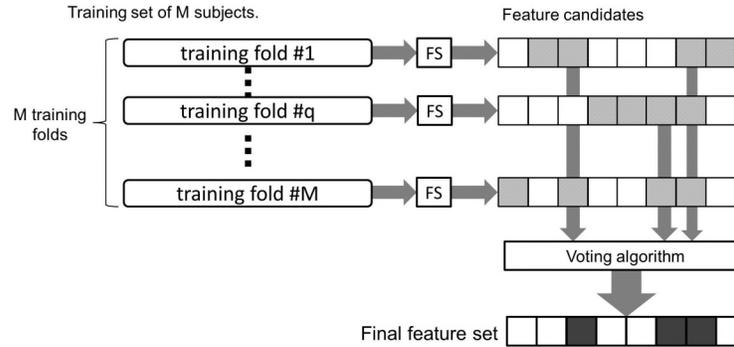


Figure 3.4: **Feature selection.** A training set of size M is subsampled M times by leaving one subject out every time. A feature selection algorithm (FS) is applied on each fold. After feature candidates are obtained, the voting algorithm reviews the agreement among the folds and produces the final feature set that it is used by the SVM.

Correlation-Based Feature Selection

Correlation-based feature selection analyzes the linear dependency between features and classification groups. Features that have a high correlation with the group labels (relevance) are good candidates for the prediction of the subject's group [15]. Similarly to previous work, our approach accounts for inter-feature correlation. Features with high inter-feature correlation (redundancy) are penalized producing a set where features tend to be linearly independent of each other [18–20].

Correlation-based feature selection was applied to each subsampled training fold in three steps: relevance evaluation, redundancy evaluation, and final scoring.

First, *feature relevance* is evaluated using a Pearson correlation coefficient. Let $N (= M - 1)$ be the total number of subjects in the current fold; $f_{i,s}$ the value of the feature i evaluated in the

subject s ; and t_s the binary group label (-1,+1) for subject s (for binary classification). Then, the relevance for feature i is obtained by:

$$\rho(i) = \frac{\left| \sum_{s=1}^N (f_{i,s} - \bar{f}_i)(t_s - \bar{t}) \right|}{\sqrt{\sum_{s=1}^N (f_{i,s} - \bar{f}_i)^2 \sum_{s=1}^N (t_s - \bar{t})^2}} \quad (3.4)$$

where \bar{f}_i and \bar{t} are the feature and the label averages respectively, evaluated using all the subjects in the current fold. After this, the top $2K$ features are selected.

In the second step, *feature redundancy* is calculated as the average of the correlation between each feature and the remaining $2K - 1$ features selected on the first step (inter-feature correlation). Features that have a high redundancy do not provide additional information due to high collinearity, and can be discarded. Let i and j be two different features ($i \neq j$), the inter-feature correlation, in the current fold, is calculated as it follows:

$$\delta(i, j) = \frac{\sum_{s=1}^N (f_{i,s} - \bar{f}_i)(f_{j,s} - \bar{f}_j)}{\sqrt{\sum_{s=1}^N (f_{i,s} - \bar{f}_i)^2 \sum_{s=1}^N (f_{j,s} - \bar{f}_j)^2}} \quad (3.5)$$

Then, the redundancy for feature i in the current fold, is calculated as:

$$\delta(i) = \frac{\sum_{j=1; j \neq i}^{2K} \delta(i, j)}{2K - 1} \quad (3.6)$$

The third step computes the score, where for each feature in the $2K$ set, the feature-label correlation (relevance) is divided by the respective average inter-feature correlation (redundancy).

The scoring function S for feature i in the current fold, was then defined as:

$$S(i) = \frac{\rho(i)}{\delta(i)} \quad (3.7)$$

Finally, the $2K$ scored features are selected as the feature set for the current fold.

ANOVA-Based Feature Selection

To the best of our knowledge, there is very little literature on the use of ANOVA as a feature selection method for classification of structural brain images. The use of ANOVA for feature selection in neuroimaging, has mainly focused on classification of fMRI datasets [21–23]. Nonetheless, ANOVA has been used in other areas of science to attenuate the *curse-of-dimensionality* [16] by discarding variables with poor statistical significance thereby reducing the size of the feature space. We explored this approach as an alternative to the more commonly used correlation-based method. Similar to the latter, the ANOVA approach is a supervised feature selection method where the *a priori* knowledge of the subject groups (ground truth) is used in the evaluation of the features. Unlike the correlation-based method introduced earlier, the ANOVA-based method does not penalize redundant features.

This method proceeds as follows: The subjects in the current training set fold (as defined above) are first divided into groups according to the classification labels of the given experiment (i.e. patients vs. controls). Then, a one-way ANOVA test is performed for each feature between the groups. The *null hypothesis* for the ANOVA test is that the mean value for each feature is the same between the groups. If there is evidence that a feature mean is significantly different between groups, then that feature becomes a good candidate for the prediction of the group. Finally, a scoring function is assigned to quantify the *relevance* of each feature according to the result of its ANOVA test.

For any feature i , with $0 < i \leq T$ where T is 936, the total number of features, ($K < T$), the ANOVA-based score S for feature i in the current fold is defined as:

$$S(i) = 1 - \alpha_i \quad (3.8)$$

where α_i is the p-value resulting of the correspondent F-test for feature i . In other words, significant features have low *p-value* and a correspondingly high score. Features with the top $2K$ scores are selected to produce the feature set for the current fold.

Voting Strategy

For both the correlation and ANOVA-based methods, the final feature set is built on a feature-by-feature basis examining the scores obtained by each feature among the participating folds.

For any given feature i , the *vote* collected from the fold q is given by:

$$V(i, q) = \begin{cases} S(i) & \text{when } i \in q \\ 0 & \text{when } i \notin q \end{cases} \quad (3.9)$$

Thus the *vote* is 0 if the feature i was not scored in fold q , and is $S(i)$, the scoring function, otherwise. The agreement is reached by:

$$A(i) = \frac{\sum_{q=1}^M V(i, q)}{M} \quad (3.10)$$

The *agreement function* $A(i)$ becomes a score average when feature i has been scored across all folds. Otherwise, infrequent features are penalized even if they have scored highly in a given particular fold. The purpose of this penalization is to enforce feature stability, producing feature sets that are robust.

Given that each voting fold scores $2K$ features, after the agreement function has been applied, only features in the top K agreement values are retained. These features become the final feature set that is used for classification.

3.2.9 Dimensionality reduction

Dimensionality reduction of high-dimensional feature spaces has been used in neuroimaging to reduce computational complexity, by projecting the high dimensional data onto a space of smaller dimensionality without loss of information [24]. Techniques such as principal component analysis (PCA) [25] have been applied directly to unfiltered highly-dimensional feature spaces in classification problems. However it has been shown that the resulting condensed space may carry over noise from original features, negatively affecting classification perfor-

mance [26].

As an alternative, we present a two-step approach, where the feature space is filtered using the proposed feature selection stage, followed by the application of PCA to the resulting space. Applying PCA to the feature space creates a new space where each dimension is *uncorrelated*. This could reduce noise and improve classification performance on this space [27].

There is, however, a trade-off when using PCA. On one hand, PCA is an *unsupervised* method. This means that class labels are not required. On the other hand, the number of dimensions or *principal components* to which the feature space is reduced to, is a parameter that must be set by the user. To take advantage of the PCA technique and tackle the selection of the number of principal components, this parameter was included in the optimization step described below.

3.2.10 SVM selection and parameter optimization

A linear SVM has only one parameter C which determines the level of allowed regularization computing the decision boundary. A low C creates a decision boundary with a large margin between the two classes, while a high C attempts to classify all the data points correctly producing a narrow margin. In contrast, a non-linear, radial basis function SVM has two parameters: C , the regularization parameter and γ which determines the weight of individual observations in the resulting decision boundary. A low γ produces higher weighting while a high γ produces lower weighting (Figure 3.5).

Clearly, the choice of SVM brings along parameters that need to be optimized. Parameter selection can either boost or hinder the generalization ability of the SVM. To determine the optimal parameters, a cross-validated *grid search* was employed on the parameter space as follows: The training set was divided into evaluation and validation partitions using cross-validation (CV) and an SVM was trained using parameters taken from the parameter grid onto the evaluation partition. The validation partition was then used to measure the accuracy of the classifier. This process was repeated for all possible training and validation partitions given by CV to obtain an average measurement of classification accuracy. This CV process was repeated

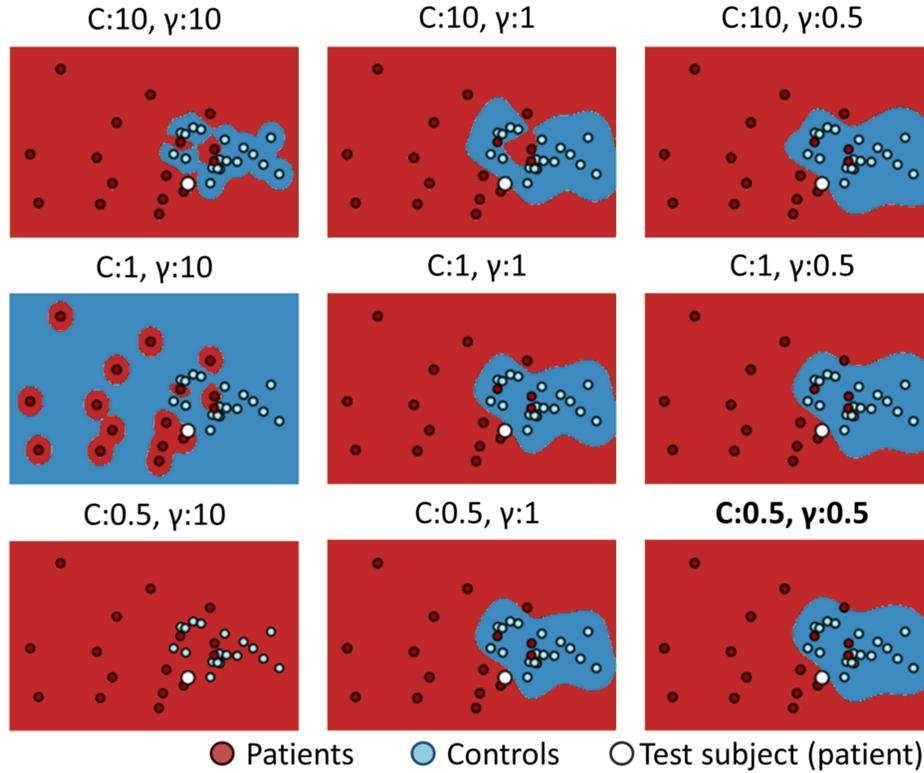


Figure 3.5: **Grid search for RBF SVM parameter selection.** Here, the SVM parameters: C (regularization) and γ (RBF Gaussian kernel width) are explored over the parameter space. The test subject is correctly identified in all cases, however each scenario presents a different decision boundary that may or not be adequate (this is the reason for parameter optimization). Overfitting is evidenced in the first column. From left to right, the decision boundary becomes more convex, generalizing better to the expense of accepting some misclassification.

for all possible parameter sets in the parameter grid. Finally the parameter set that produced the best average accuracy was selected as the optimal choice to configure the SVM (Figure 3.6).

If the model being evaluated included the PCA reduction step, a range for the number of principal components was included into the parameter grid. In that case, parameter optimization occurred on the PCA-reduced space.

3.3 Results

Each of the eight classification models was evaluated on the full range of possible K values (from 2 to 936 features for the universal feature space (T1+T2+FA+MD), and from 2 to 234

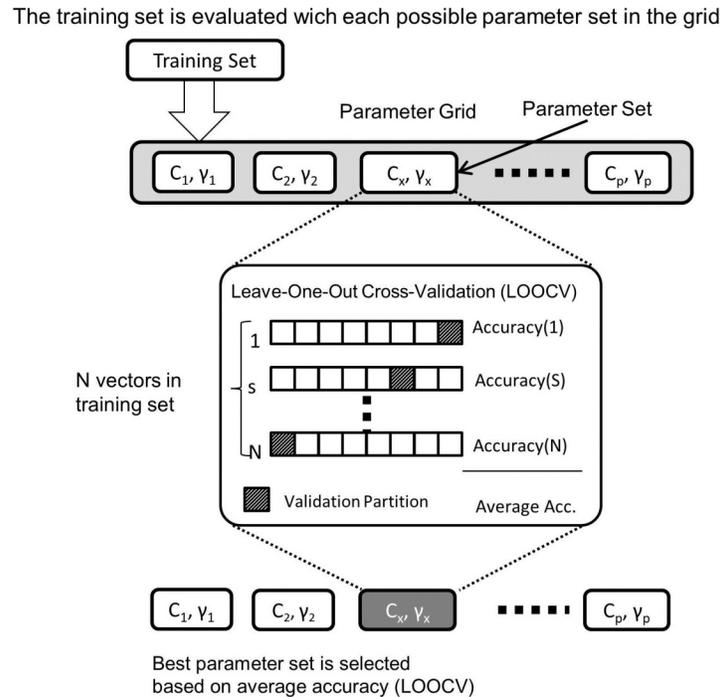


Figure 3.6: **Parameter Selection.** A classifier is trained using a training set and one parameter set from the parameter grid. The training set is the same for all classifier instances while the parameter set varies with the goal of evaluating the best parameter configuration for a given type of classifier and training set. A leave-one-out cross-validation procedure is performed on the training set to evaluate the accuracy of the classifier for a given parameter configuration. The best average accuracy determines the set of parameters that will be used. C (regularization) , γ (RBF Gaussian kernel width)

for individual image subspaces). Given that K is the only parameter that needs to be set by the user, an adequate range for the choice of K is addressed in the discussion section.

3.3.1 Detection of TLE patients

A maximum accuracy of 88.9% for automatic classification between patients and controls was obtained by an *anova-pca-svm-linear* model using 10 features from the T1 image only as shown in Table 3.1. This model misclassified 3 patients, all of whom were diagnosed with R-TLE, and one control volunteer. In general the classification models that employed features from the T1 image, obtained 81% average accuracy, followed by the MD models with an average accuracy of 75%, T2 models with an average accuracy of 74% and in last place FA models with an accuracy of 67%.

Table 3.1: **Classification results for Experiment I - TLE detection.** The best performance for each model is shown. Each column indicates the image from where features were selected. T1: quantitative T1 image, T2: quantitative T2 image, FA: fractional anisotropy image, MD: mean diffusivity image.

Configuration	T1	T2	FA	MD	ALL
<i>correlation-svm-linear</i>					
Accuracy	77.8%	77.8%	69.4%	72.2%	80.6%
Sensitivity	58.8%	76.5%	64.7%	58.8%	64.7%
Specificity	94.7%	78.9%	73.7%	84.2%	94.7%
Cardinality	4	5	22	5	170
<i>correlation-pca-svm-linear</i>					
Accuracy	86.1%	72.2%	72.2%	80.6%	77.8%
Sensitivity	88.2%	76.5%	64.7%	64.7%	64.7%
Specificity	84.2%	68.4%	78.9%	94.7%	89.5%
Cardinality	14	5	47	106	14
<i>correlation-svm-rbf</i>					
Accuracy	77.8%	77.8%	52.8%	77.8%	66.7%
Sensitivity	82.4%	64.7%	0%	82.4%	70.6%
Specificity	73.7%	89.5%	100%	73.7%	63.2%
Cardinality	9	2	14	10	8
<i>correlation-pca-svm-rbf</i>					
Accuracy	80.6%	77.8%	75.0%	80.6%	77.8%
Sensitivity	82.4%	64.7%	76.5%	82.4%	70.6%
Specificity	78.9%	89.5%	73.7%	78.9%	84.2%
Cardinality	6	2	47	59	17
<i>anova-svm-linear</i>					
Accuracy	77.8%	72.2%	72.2%	69.4%	80.6%
Sensitivity	76.5%	64.7%	76.5%	58.8%	64.7%
Specificity	78.9%	78.9%	68.4%	78.9%	94.7%
Cardinality	51	21	37	8	530
<i>anova-pca-svm-linear</i>					
Accuracy	88.9%	66.7%	69.4%	75.0%	75.0%
Sensitivity	82.4%	58.8%	58.8%	58.8%	64.7%
Specificity	94.7%	73.7%	78.9%	89.5%	84.2%
Cardinality	10	12	27	8	13
<i>anova-svm-rbf</i>					
Accuracy	75.0%	77.8%	55.6%	72.2%	63.9%
Sensitivity	88.2%	88.2%	52.9%	58.8%	76.5%
Specificity	63.2%	68.4%	57.9%	84.2%	52.6%
Cardinality	7	7	3	7	12
<i>anova-pca-svm-rbf</i>					
Accuracy	80.6%	72.2%	66.7%	75.0%	77.8%
Sensitivity	82.4%	70.6%	58.8%	58.8%	70.6%
Specificity	78.9%	73.7%	73.7%	89.5%	84.2%
Cardinality	11	7	18	8	16

Figures 3.7 and 3.8 (I) show regions selected by the best classification model that were common to all subjects. Figure 3.7 shows left/right mean intensity features while Figure 3.8 shows regions selected based on asymmetry. Thus, the middle-temporal cortex, the right parahippocampal white matter and the left entorhinal white matter are selected based on their mean intensity. In contrast, the inferior-temporal, middle-temporal and parahippocampal white matter regions are selected based on their asymmetry. These findings are consistent with previous studies where changes in temporal white matter regions have been associated with TLE [28, 29]. Although feature selection is not restricted to the temporal lobe, it is relevant to

notice that all the regions shown in the overlay are temporal lobe regions.

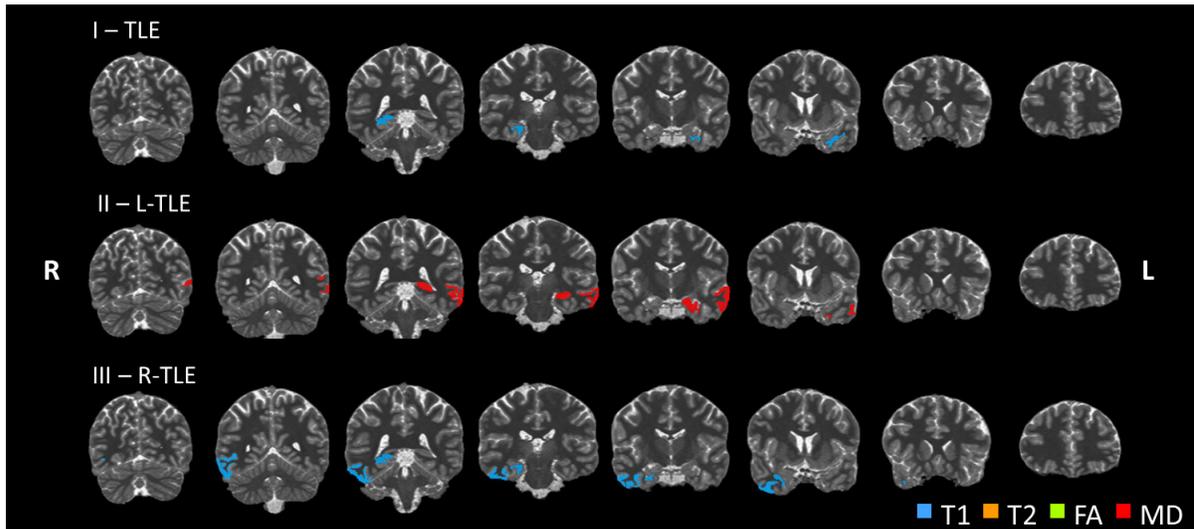


Figure 3.7: **Stable ROI selected based on left/right mean intensity.** The best classification cases across subjects are shown. The method used only features in the left temporal cortex to identify L-TLE patients. Similarly, cortical regions in the right temporal lobe were identified as key for classifying R-TLE patients. I) TLE detection: [anova-pca-svm-linear, T1, $K = 10$]. II) L-TLE detection: [correlation-pca-svm-rbf, MD, $K = 7$]. III) R-TLE detection: [correlation-pca-svm-linear, T1, $K = 29$].

We hypothesized that the heterogeneity of the patient group (L-TLE, R-TLE) was a relevant factor that influenced the accuracy of the patient identification process. Its effect can be seen by performing a dimensionality reduction transformation on the feature space (Figure 3.9). While the control group forms a cluster, the patient groups are sparse. This configuration was consistently reproducible for feature spaces of different sizes and for different data projection techniques including PCA, Isomaps and Multidimensional Scaling. These projections demonstrated that the L-TLE group tends to be linearly separable from the control group, while the R-TLE group intersected with the control group. This configuration could explain why perfect classification is achievable between controls and L-TLE, while classification between controls and R-TLE appeared to be more challenging. To address the heterogeneity matter, classification accuracy was examined using more homogeneous patient subgroups: L-TLE vs. controls (experiment II) and R-TLE vs. controls (experiment III).

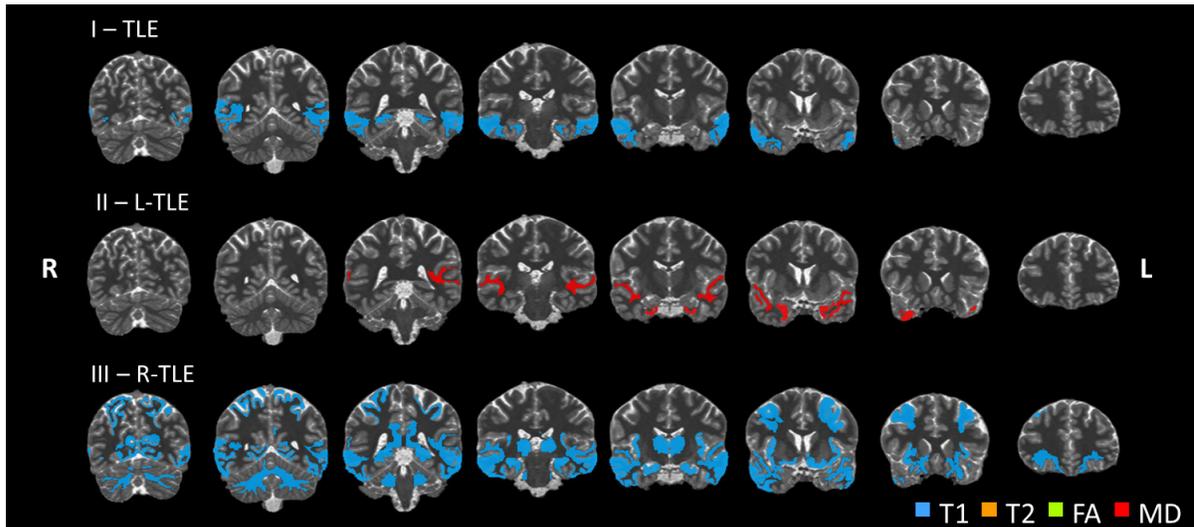


Figure 3.8: **Stable ROI selected based on asymmetry.** The best classification cases across subjects are shown. L-TLE patients are successfully classified by looking at the asymmetry in the temporal white matter regions whereas R-TLE patients are harder to classify driving the method to look outside the temporal lobe. I) TLE detection: [anova-pca-svm-linear, T1, $K = 10$]. II) L-TLE detection: [correlation-pca-svm-rbf, MD, $K = 7$]. III) R-TLE detection: [correlation-pca-svm-linear, T1, $K = 29$].

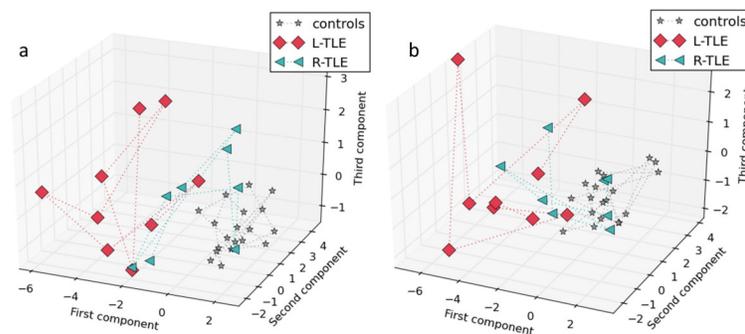


Figure 3.9: **Feature space projection.** These tridimensional projections are obtained by performing a PCA transformation from an original feature space with 14 dimensions ($K=14$) for experiment I. a) using correlation-based feature selection b) using ANOVA-based feature selection. In both projections the control group tends to form a cluster while the patient subgroups are sparse. L-TLE: left TLE, R-TLE: right TLE

3.3.2 Detection of L-TLE patients

Perfect classification (100% accuracy) was attainable for L-TLE patients (Table 3.2). This result was reported by the *correlation-pca-svm-rbf* model with MD features only (K=7). However, the *correlation-pca-svm-linear* and *anova-pca-svm-linear* models also obtained good classification accuracy (96.4%) using MD features with low feature set cardinalities. In general, linear models were not outperformed by non-linear models. Overall, PCA improved classification accuracy.

Table 3.2: **Classification results for Experiment II - Detection of L-TLE patients.** The best performance for each model is shown. The best classification spaces were MD and T1. In general, linear models are not outperformed by their non-linear versions. Overall, PCA improves classification accuracy.

Configuration	T1	T2	FA	MD	ALL
<i>correlation-svm-linear</i>					
Accuracy	92.9%	78.6%	85.7%	96.4%	89.3%
Sensitivity	100%	55.6%	66.7%	88.9%	66.7%
Specificity	89.5%	89.5%	94.7%	100%	100%
Cardinality	7	18	5	7	6
<i>correlation-pca-svm-linear</i>					
Accuracy	92.9%	71.4%	85.7%	96.4%	92.9%
Sensitivity	100%	44.4%	66.7%	88.9%	77.8%
Specificity	89.5%	84.2%	94.7%	100%	100%
Cardinality	4	58	47	40	138
<i>correlation-svm-rbf</i>					
Accuracy	85.7%	67.9%	89.3%	92.9%	92.9%
Sensitivity	77.8%	0%	66.7%	77.8%	77.8%
Specificity	89.5%	100%	100%	100%	100%
Cardinality	4	8	5	5	6
<i>correlation-pca-svm-rbf</i>					
Accuracy	96.4%	75.0%	85.7%	100%	92.9%
Sensitivity	100%	22.2%	55.6%	100%	77.8%
Specificity	94.7%	100%	100%	100%	100%
Cardinality	4	96	3	7	23
<i>anova-svm-linear</i>					
Accuracy	85.7%	78.6%	85.7%	92.9%	89.3%
Sensitivity	88.9%	55.6%	66.7%	88.9%	77.8%
Specificity	84.2%	89.5%	94.7%	94.7%	94.7%
Cardinality	5	41	33	10	9
<i>anova-pca-svm-linear</i>					
Accuracy	92.9%	71.4%	85.7%	96.4%	92.9%
Sensitivity	88.9%	44.4%	66.7%	88.9%	77.8%
Specificity	94.7%	84.2%	94.7%	100%	100%
Cardinality	5	66	16	12	10
<i>anova-svm-rbf</i>					
Accuracy	85.7%	67.9%	75.0%	78.6%	78.6%
Sensitivity	66.7%	0%	22.2%	55.6%	55.6%
Specificity	94.7%	100%	100%	89.5%	89.5%
Cardinality	5	8	6	4	4
<i>anova-pca-svm-rbf</i>					
Accuracy	92.9%	71.4%	82.1%	96.4%	92.9%
Sensitivity	77.8%	22.2%	44.4%	88.9%	77.8%
Specificity	100%	94.7%	100%	100%	100%
Cardinality	4	18	7	17	22

When looking at individual feature subspaces (columns on Table 3.2), the best accuracies across models were obtained by the MD and T1 subspaces where the average accuracy was 90% for T1, and 94% for MD. The average classification accuracy across models dropped in the FA and T2 subspaces to 84% and 73% respectively. In general, classifiers had an excellent performance on the universal feature space where the average accuracy was 90%.

Figures 3.7 and 3.8 (II) show the regions selected by the best classification method that were common to all subjects. In general, the asymmetry between left and right temporal lobe white matter regions, and the mean intensity of left temporal cortical regions are determined to be relevant. Specifically, the left hippocampus, the left middle-temporal cortex and the left entorhinal white matter regions are selected based on the *mean intensity*, whereas the entorhinal cortex, the superiotemporal white matter and the temporal pole are selected based on their *asymmetry*. All of these regions clinically correlate to pathologies in L-TLE patients such as mesial temporal sclerosis (MTS) and focal cortical dysplasia (FCD).

3.3.3 Detection of R-TLE patients

As presumed by the PCA projection, the classification between R-TLE patients and controls was a difficult problem. This was reflected in classification accuracy average of 74% across the different models as well as the high cardinality reported by each model (Table 3.3).

A possible explanation for these results is that the feature space does not have discriminative features to distinguish the R-TLE patients from controls (evidenced by the PCA projection). Hence, reliable decision boundaries cannot be obtained. There is some evidence supporting the proximity of R-TLE patients and controls in this type of images. For example, Zhong Xue et al. [28] have found fewer regions to distinguish R-TLE patients from controls than L-TLE from controls. Similarly, Ahmadi et al. [30] have reported fewer and less extensive gray matter change in R-TLE than in L-TLE patients with respect to controls using voxel-based morphometry.

The average accuracy for T1 and T2 subspaces across models was 74% and 73% respectively. For FA and MD, it was 76% and 74%. Similarly to the classification of L-TLE patients,

Table 3.3: **Classification results for Experiment III - Detection of R-TLE patients.** The best performance for each model is shown. The classification of R-TLE patients is a difficult problem in the proposed feature space. This is reflected by the high cardinality of the optimal solutions as well as the average accuracy of 73% across models.

Configuration	T1	T2	FA	MD	ALL
<i>correlation-svm-linear</i>					
Accuracy	77.8%	77.8%	81.5%	74.1%	88.9%
Sensitivity	37.5%	37.5%	62.5%	37.5%	62.5%
Specificity	94.7%	94.7%	89.5%	89.5%	100%
Cardinality	12	143	52	34	141
<i>correlation-pca-svm-linear</i>					
Accuracy	81.5%	77.8%	81.5%	77.8%	70.4%
Sensitivity	62.5%	37.5%	50%	37.5%	37.5%
Specificity	89.5%	94.7%	94.7%	94.7%	84.2%
Cardinality	29	21	13	56	52
<i>correlation-svm-rbf</i>					
Accuracy	66.7%	63%	66.7%	59.3%	59.3%
Sensitivity	25%	25%	25%	12.5%	25%
Specificity	84.2%	78.9%	84.2%	78.9%	73.3%
Cardinality	8	8	7	7	6
<i>correlation-pca-svm-rbf</i>					
Accuracy	74.1%	74.1%	77.8%	77.8%	74.1%
Sensitivity	25.0%	12.5%	50%	25.0%	12.5%
Specificity	94.7%	100%	89.5%	100%	100%
Cardinality	11	53	13	83	179
<i>anova-svm-linear</i>					
Accuracy	77.8%	74.1%	81.5%	74.1%	77.8%
Sensitivity	37.5%	50%	50%	37.5%	25.0%
Specificity	94.7%	84.2%	94.7%	89.5%	100%
Cardinality	56	9	37	110	156
<i>anova-pca-svm-linear</i>					
Accuracy	77.8%	77.8%	81.5%	81.5%	70.4%
Sensitivity	50%	37.5%	50%	50%	25.0%
Specificity	89.5%	94.7%	94.7%	94.7%	89.5%
Cardinality	29	43	122	66	30
<i>anova-svm-rbf</i>					
Accuracy	63%	66.7%	63%	66.7%	55.6%
Sensitivity	25%	12.5%	12.5%	12.5%	25%
Specificity	78.9%	89.5%	84.2%	89.5%	68.4%
Cardinality	7	9	7	6	8
<i>anova-pca-svm-rbf</i>					
Accuracy	74.1%	70.4%	77.8%	77.8%	77.8%
Sensitivity	37.5%	37.5%	25.0%	25.0%	37.5%
Specificity	89.5%	84.2%	100%	100%	94.7%
Cardinality	7	25	78	12	15

the best accuracy on individual image spaces was obtained on the MD space. The highest accuracy on the universal feature space was 88.9% and it was obtained using a *correlation-svm-linear* model with 141 features. Intuitively, models with such a high number of features are less useful to the researcher than simpler models. One of such alternatives is presented by the *correlation-pca-svm-linear* model on T1 with $K=29$ and an accuracy of 81.5%.

Figures 3.7 and 3.8 (III) show the regions that were common to all subjects, reported by the aforementioned simpler T1 model. Regions selected by virtue of their *mean intensity* are shown in Figure 3.7. These regions are the right inferiotemporal cortex, the right entorhinal and parahippocampal white matter regions as well as the right white matter in temporal pole. Neither extra-temporal nor left-temporal regions were chosen.

When looking at regions selected based on *asymmetry*, Figure 3.8 reveals numerous features most of them corresponding to the temporal lobes, including the inferiotemporal cortex, the parahippocampal white matter and in general the temporal lobe white matter. The difficulty of the classification between R-TLE and controls explains why the optimal results are obtained by the inclusion of extra-temporal features. To validate this assumption we restricted the method to select only temporal lobe features and we found that the classification accuracy reduced from 81.5% to 70.4%.

3.3.4 Evaluation of feature selection methods

After obtaining the classification results, we performed an analysis of the two feature selection strategies. The two methods were compared measuring their *stability* and *similarity*. To measure the stability of each method we calculated the *average Tanimoto distance* among feature sets [31, 32] in the outer leave-one-out cross-validation loop (training/testing loop). Figure 3.10 shows the results. As expected, stability increases asymptotically towards 100% as the number of requested features approaches the number of available features. The $[2 < K < 100]$ range of cardinalities is most interesting as stability varies between 40% and 70% (right column in the figure). This range gives a better estimation of stability as the methods are asked to retrieve at most 100 features out of 936 *for each iteration* of the cross-validation loop. In comparison, in

the absence of any procedure, the likelihood of selecting 100 features out of 936 for the same 36 subjects is approximately 10^{-36} . As shown in the figure, the most stable features ($> 80\%$) were obtained in experiment II (L-TLE) and the least stable features ($< 70\%$) in experiment III (R-TLE). In all three experiments it was observed that classification accuracy was proportional to feature stability.

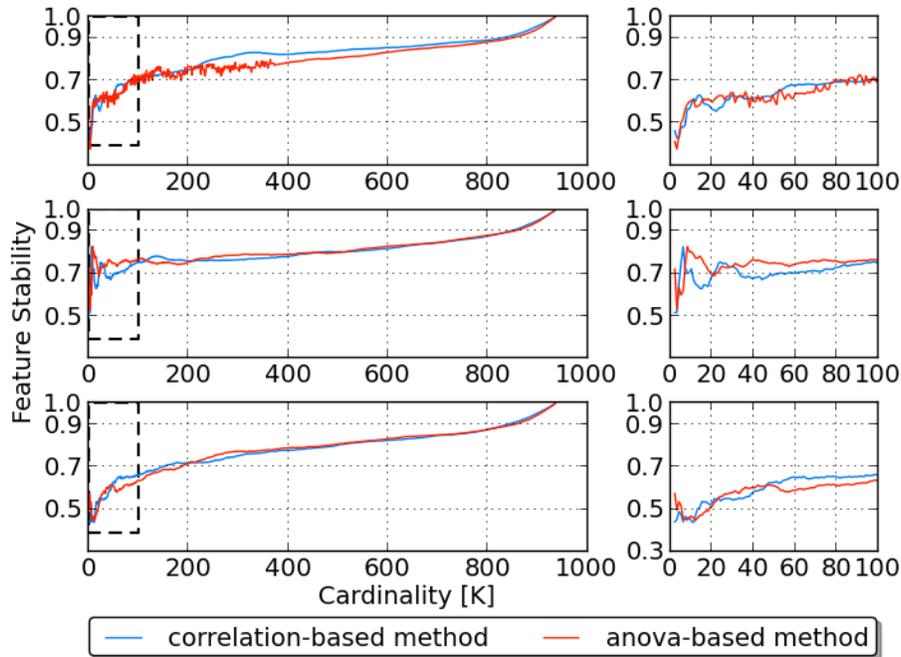


Figure 3.10: **Feature stability.** Each training set obtained by leave-one-out cross-validation produces one feature set. The average Tanimoto distance among these sets gives an indication of stability. a) TLE detection b) L-TLE identification c) R-TLE identification. The right column shows a zoomed-in version of the dotted rectangle on the left.

To assess similarity between methods, in a given training fold, the Tanimoto distance between the ANOVA-generated feature set and the correlation-generated feature set was measured. This was repeated for all the training folds generated by the LOOCV procedure. Then, the average Tanimoto distance was reported. Figure 3.11 shows how similarity varies according to cardinality. For very small cardinalities ($K < 20$) the two methods (correlation, ANOVA) generate very similar features. In the $[20 < K < 100]$ range (right column in the figure) the similarity drops to 50%. This is desirable as it shows that the two methods are sufficiently distinct

in the cardinality range that contains both the best classification results and the inherent dimensionality of the problem (This is further addressed in the Discussion, below). Conceptually, the dissimilarity between the two methods can be explained by the redundancy evaluation step in the correlation method. However, as the cardinality approaches the number of available features, the correlation method runs out of features to discard and the similarity between methods increases.

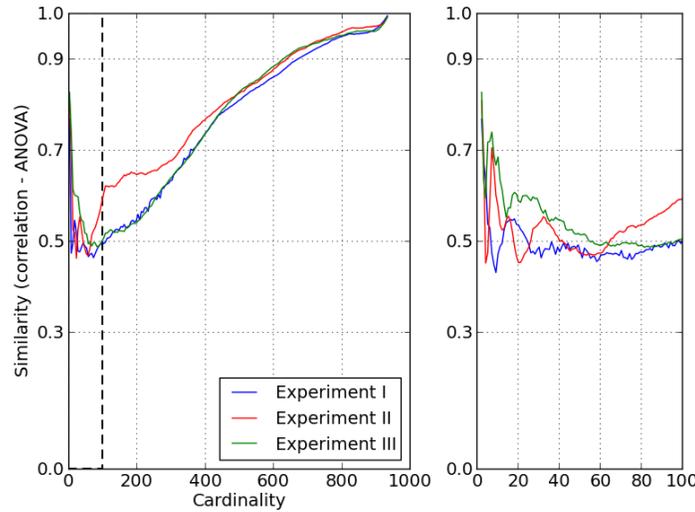


Figure 3.11: **Similarity between the correlation-based and the ANOVA-based feature selection methods.** The right column shows the region outlined by the dashed rectangle on the left.

3.3.5 Informative features for TLE detection

Informative features were evaluated by analyzing how frequently a feature was selected when the classifier was successful. The level of success was defined by setting a minimum accuracy threshold, and the analysis was restricted to those cardinalities where the classifier performance was higher than the threshold. For each cardinality the feature frequency was measured on the external LOOCV (training/testing) analyzing the features obtained from each training fold. Then, a ranking was obtained by averaging frequencies across the selected cardinalities. The regions of interest corresponding to the top ranked features are analyzed in the Discussion.

Alternatively, we considered using the coefficient assigned to features by the linear clas-

sifiers as a measurement of their relevance. However this approach would not reflect the importance of the features in non-linear classifiers or in classifiers that use PCA. We think that our approach is more comprehensive as it identifies features that are relevant and stable across cardinalities regardless of the type of classification model.

3.3.6 Effect of the sample size on the reliability of the results

We quantified the effect of the *small sample size* in the performance of the classifier by evaluating the *SVM reliability index (SRI)* [33]. Let \mathbf{w} be the vector defining the decision boundary (vector normal to the decision plane), this vector is obtained as the solution of the respective convex optimization problem. Let \mathbf{w}^* be the alternative convex optimization solution after randomly removing some data points. If we have *enough data for training* then we can randomly remove some and what is left will result in $\mathbf{w}^* \approx \mathbf{w}$. If we do not have enough data, the random removal of training data will result in a very different decision boundary and $\mathbf{w}^* \neq \mathbf{w}$. This is quantified by the SRI as follows:

$$SRI(\mathbf{w}^*, \mathbf{w}) = |r(\mathbf{w}^*, \mathbf{w})| \quad (3.11)$$

which is the absolute value of the Pearson product-moment correlation coefficient between \mathbf{w}^* and \mathbf{w} . The SRI was evaluated for each experiment by selecting a training set to estimate \mathbf{w} , then the training set was randomly subsampled 10 times to obtain \mathbf{w}^* estimates. Each SRI result was averaged. This process was repeated for all training sets.

It is expected that the decision boundary remains stable when data points that do not weight in its calculation are removed. However, when support vectors are eliminated, this causes the boundary to be redefined and consequently the SRI decreases. Our results show that the decision boundary is fairly sensitive to the training set size (Figure 3.12). However, in all three experiments we obtained SRI measures above 80% (arbitrarily set).

Notice that the stability of the boundary is independent from performance: if two classes are sufficiently distinct (large margin), then they can be linearly separable by *many* possible

boundaries. From the PCA projections and the experimental results, we believe this is the case for the L-TLE group vs. controls.

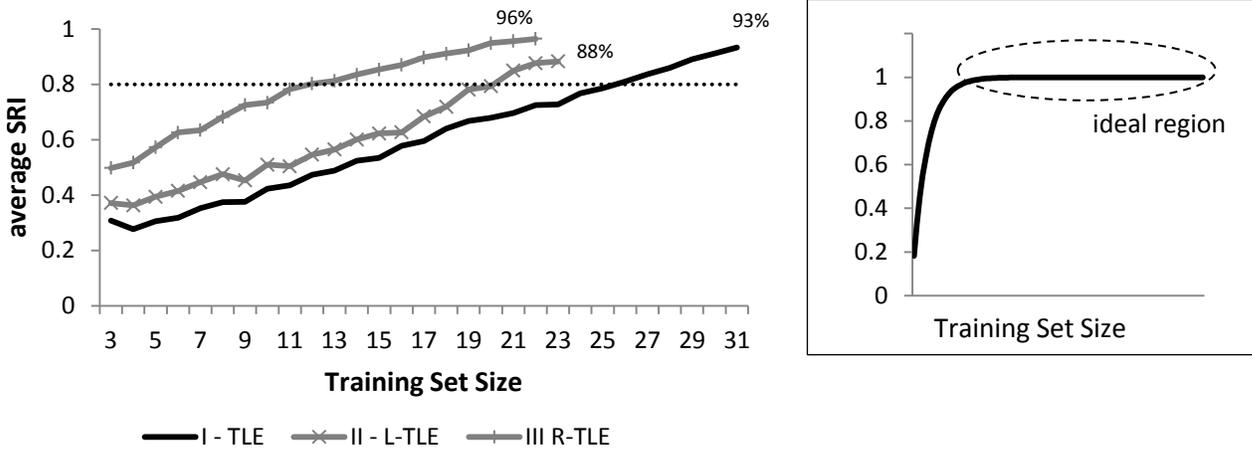


Figure 3.12: **SVM reliability index (SRI)**. In all three experiments the reliability of the boundary was higher than 80%. A larger dataset would improve the robustness of the boundary with respect to the training set size.

3.4 Discussion

3.4.1 Analysis of informative features

Table 3.4 summarizes the ROIs associated with informative features, and Figure 3.13 this information is grouped by lobe, type of feature, and image type. In all three experiments most of these region belong to the temporal lobe (a total of 71 features in the table). Among all, the middle-temporal, superior-temporal, the temporal pole, and hippocampal ROIs are common to all three classification experiments revealing the importance of these regions for the identification of patients. The selection of the hippocampus is to be expected, since it is the presumed focus in mesial temporal sclerosis (MTS), the most common pathology in this group of patients [34]. Informative features in the neocortex and adjacent white matter could relate either to changes due to seizure propagation, such as gliosis or neuronal loss, or could be related to the presence of epileptogenic lesions in the neocortex, considering that the patient group also included subjects with focal cortical dysplasia (FCD). The inclusion of some extra-temporal

regions (in the frontal lobe and parietal lobes, as well as the occipital lobe) is also notable, however, here the features likely relate to white matter abnormalities associated with seizure propagation, and not extra-temporal epileptogenic lesions, since extra-temporal onset was not observed clinically in these patients.

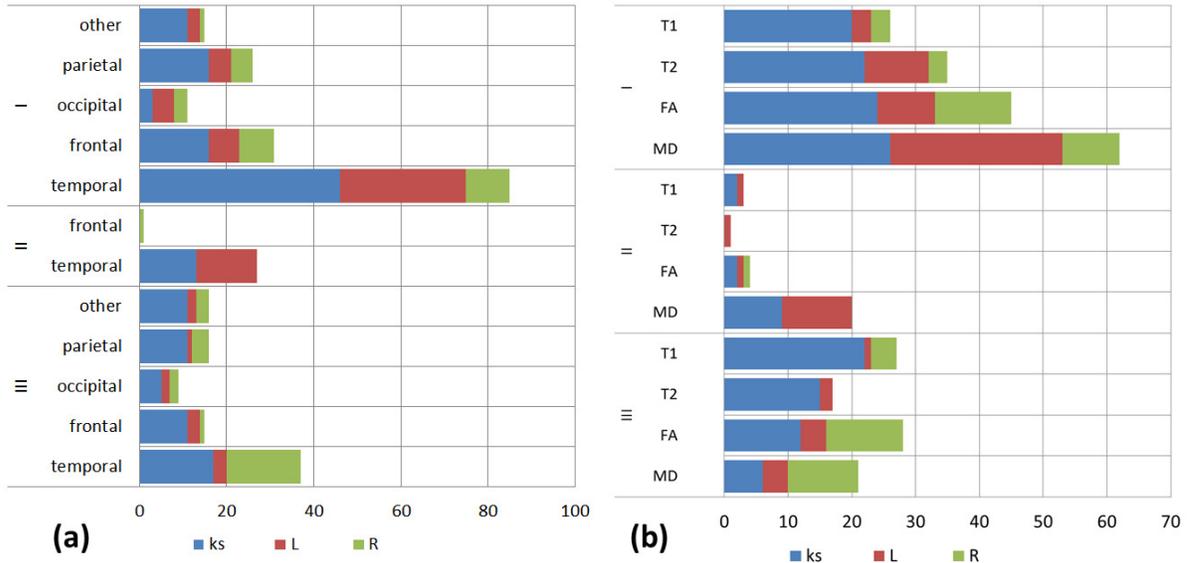


Figure 3.13: **Relevant features for classification by kind.** (a) by lobe, (b) by image type. In all three experiments most of the features originated in the temporal lobe. Across all experiments, asymmetry (ks) plays an important role in classification. Mean intensity features in the left hemisphere (L) were preferred over mean intensity features in the right hemisphere (R) to classify L-TLE patients (II). The opposite case was evidenced for the classification of R-TLE patients (III).

We also analyzed relevant features based on the type of feature. We found that asymmetry features (ks) were chosen more often than mean intensity features from each hemisphere, highlighting the benefit of sensitive examination of intensity distributions between bilateral regions. In patients with unilateral TLE was expected that asymmetry features were highly relevant since seizure onset is restricted to one hemisphere. However, asymmetry features may also be sensitive to compensatory mechanisms occurring in the contralateral hemisphere and thus may not be specific to seizure-related abnormalities. When examining individual image spaces, the classifiers performed optimally in the T1 and MD subspaces, and generally good in FA and T2. Also, classifiers on the universal space obtained accuracies that were better or close to the best individual image subspace. We also see that in experiments I and II, MD

Table 3.4: **Summary of ROI identified by informative features.** The values reported are the number of informative features referring to the respective region. Most of the features indicate the relevance of temporal lobe regions in the classification experiments (71 in total). The enclosed regions are shared by all three classification experiments.

Lobe / ROI	EXPERIMENT				Lobe / ROI	EXPERIMENT			
	I-TLE	II-L-TLE	III-RTLE	Total		I-TLE	II-L-TLE	III-RTLE	Total
temporal	44	5	22	71	frontal	18		9	27
bankssts	5		2	7	caudal anterior cingulate	1			1
entorhinal	5	1		6	caudal middle-frontal	2		2	4
fusiform	5		3	8	frontal pole	2		1	3
Hippocampus	2	1	2	5	lateral orbito-frontal	1		2	3
inferior-temporal	8		1	9	medial orbito-frontal	6		2	8
middle-temporal	5	1	5	11	pars opercularis	3			3
para-hippocampal			3	3	pars triangularis	1			1
superior-temporal	7	1	2	10	pre-central			1	1
temporal pole	6	1	3	10	superior-frontal	2		1	3
transverse-temporal	1		1	2	other	10		7	17
occipital	4		7	11	Accumbens area	1		3	4
cuneus	1			1	Caudate			1	1
lateral-occipital	3		3	6	Cerebellum	4		2	6
lingual			1	1	insula	2			2
pericalcarine			3	3	Putamen	1			1
parietal	11		9	20	Thalamus-Proper	2		1	3
inferior-parietal	1			1					
isthmuscingulate	1		2	3					
post-central	1		1	2					
precuneus	1		1	2					
superior-parietal	5		2	7					
supramarginal	2		3	5					

features were chosen most, followed by FA. The lateralization ability of diffusion metrics has been shown before [28–30], including in our previous work, [35], where DTI and relaxometry quantitative imaging parameters were compared in the temporal lobe. This work extends this previous findings showing how these quantitative imaging parameters across the entire brain can be used to classify patient groups.

3.4.2 Criteria for the selection of the feature set cardinality

The proposed method requires that the feature set cardinality K is set *a priori*. To determine an adequate range for K we plotted K against classification accuracy. For this analysis the best model for each experiment was selected. Due to the procedure for feature selection, as described in this chapter, relevant features are selected *first*, then, as expected, the accuracy degrades as noisy/non-relevant features are added. Figure 3.14 shows that above 200 features the average classification approaches the line of random chance. The classifiers present an

acceptable behavior with several peaks above the 80% classification rate for $K < 50$. Given that these plots are taken on the best models we used $K = 50$ as an upper bound for the selection of K .

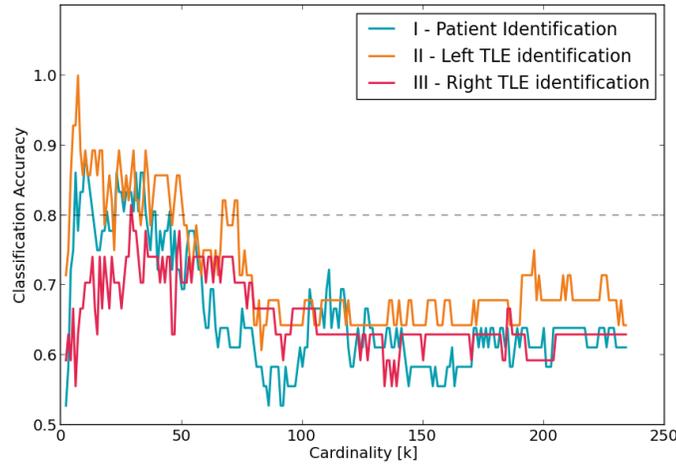


Figure 3.14: **Classification accuracy versus number of features (K parameter)**. As the requested cardinality increases, the classification accuracy drops. This is expected given that important features are picked first and adding irrelevant features will not improve accuracy but rather will add noise and increase the dimensionality of the feature space beyond the intrinsic dimension of the classification problem.

Independently, a cross-validated L1-penalized logistic regression model was used to estimate K for experiments I, II and III. The number of non-zero coefficients in the regularized regression (regularization factor = 0.1) corresponds to the estimated size of the feature set K . The results are presented in Table 3.5. In all cases the estimated K is lower than the upper bound previously obtained by simple inspection ($K < 50$). Also, it is important to notice that the estimated K (as shown in Table 3.5) is close to the number of observations (36 subjects).

Table 3.5: Estimation of K using a L1-penalized logistic regression model (regularization =0.1)

Experiment	estimated K
I - TLE detection	25
II - L-TLE detection	28
III - R-TLE detection	34

3.4.3 Criteria for the selection of a classification model

Considering classification accuracy, correlation-based models performed slightly better than those based upon ANOVA. This could be explained by the fact that the ANOVA-based method does not eliminate superfluous (collinear) features. Though in theory feature redundancy should not affect classification accuracy, the presence of redundant features in the feature set could hinder the algorithm from finding an optimal set before the desired cardinality is reached.

In terms of the type of SVM, we found that linear classifiers exhibited slightly better performance than non-linear ones. However this could be attributable to the size of our sample, which may not be large enough to adequately estimate non-linear SVM parameters.

In terms of computation time the ANOVA-based classifiers outperformed the correlation-based classifiers for large cardinalities (Figure 3.15). This estimation was performed on a machine with 4 CPU cores (Intel Core I7-2600 CPU @ 3.4GHZ) running Ubuntu Linux 12.04 with 16GB of RAM. The ANOVA-based method can evaluate several features simultaneously and it does not incur in the computational cost of the feature redundancy evaluation. Additionally, the PCA transformation and the parameter optimization for non-linear classifier increased significantly the time required to fully classify a dataset.

With these considerations in mind, a good compromise between classification accuracy and computation time can be achieved by using an *anova-pca-linear* model. A linear SVM requires less parameter tuning than a non-linear SVM and classification accuracy did not degrade as rapidly as it did for non-linear classifiers as the cardinality increased. The ANOVA approach is more scalable than the correlation-based method in terms of computational cost when evaluating large feature spaces. In addition, the PCA transformation has the effect of decorrelating the resulting feature space which explains why the classification accuracy is similar to those classifiers using the correlation-based method where redundant features are discarded.

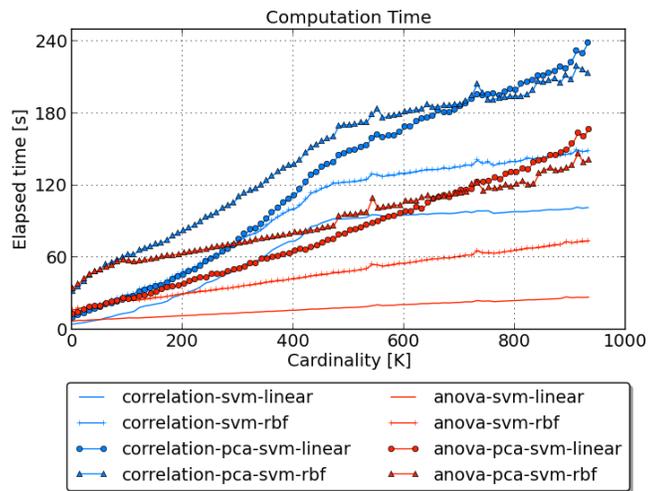


Figure 3.15: **Computational performance.** Performance comparison in terms of computation time for the eight classification models discussed in this chapter. The evaluation was made on experiment I, the largest dataset on a machine with 4 CPU cores (Intel Core I7-2600 CPU @ 3.4GHZ) running Ubuntu Linux 12.04 with 16GB of RAM.

3.4.4 Comparison with similar studies

Our results were comparable and in some cases better than to those obtained by Focke et al [1]. For example, the classification accuracy for L-TLE described by their method varied between 93% and 95%, while our method obtained perfect identification. In contrast, their accuracy for R-TLE detection was 97% while ours was 88.9%. We believe that the size of the patient group is a key factor to obtain higher accuracies in the R-TLE group. Their patient group was more than twice the size of ours (38 vs. 17). In addition, we did not perform any morphometry-based features (such as gray matter and white matter probability maps); these features provided the best performance for their R-TLE group. The purpose of this chapter was to explore classification using quantitative imaging (T1 maps, T2 maps, DTI). We plan to investigate the added benefit of adding Morphometry to these features as we believe that this could improve our results. Otherwise, in agreement with these findings, we observed that MD is more useful for discrimination of TLE patients than FA, and we also observed poor classification results when looking at T2 data only as they did. Similarly, Keihaninejad et al. [2] stated an accuracy of 86% for the identification of TLE patients where ours is 88.9%,

with the caveat that this number refers to MR-negative patients (MRI without clinical findings), therefore more difficult to classify.

3.4.5 Limitations of this work

It is well known that SVM performance is highly dependent on the quality of the training set. This is a common concern when dealing with biological data characterized by a large number of features and a small number of observations like ours. We have followed several steps to address this limitation and mitigate its effect.

On the one hand, it is clear that the number of features has a direct impact on classification error [36]. Feature sets with cardinalities above the *intrinsic dimension* of the problem can lead to overfitting effects because the classifier can produce decision boundaries that follow the sample points too closely [37]. This effect is evident in our dataset. Figure 3.14 shows that the generalization ability of the classifier suffers as the cardinality of the feature set increases. A simple method to estimate the ideal cardinality is provided (3.4.2) and the experiments corroborated optimal results in regions around this estimation.

On the other hand, a small sample size makes necessary the systematic use of cross-validation to avoid overfitting. It is well known that cross-validation reduces the error bias at the expense of increasing the error variance. We believe that we can obtain a better bias-variance trade-off by acquiring more data and reducing the amount of cross-validation. Also, a larger dataset can contribute to SVM parameter optimization, boosting the comparison between linear and non-linear models.

Another limitation in our work is the sensitivity for the detection of R-TLE patients. A larger dataset could potentially improve feature selection leading to better classification rates. Nonetheless, there is evidence suggesting that the R-TLE group is very heterogeneous in terms of MRI abnormalities [38, 39] and the low sensitivity of our results could be associated with that heterogeneity. To validate this, a larger dataset would allow the R-TLE group to be subdivided into specific pathologies (i.e. MTS, non-MTS, FCD) for classification analyses.

3.5 Conclusion

This chapter describes a novel approach for the detection of TLE patients using feature selection and support vector machine methods. The novelty of this work consists of the use of multi-parametric quantitative imaging, the definition of a measurement of regional asymmetry using a non-parametric statistical test and the evaluation and validation of the optimal cardinality of the problem. The main contribution of this chapter comprises the evaluation of key factors influencing classification performance, namely the feature selection method, the possibility of a further dimensionality reduction step with the use of PCA, and the type of support vector machine for this type of data. Then, relevant features are identified and their clinical significance is addressed.

Our results demonstrate that the identification of TLE subjects based on quantitative MR images is possible. In particular, DTI derived features seemed to be more effective than T2 features, which in general performed sub-optimally. In all experiments, good accuracies were attainable. However, the identification of R-TLE patients proved to be a difficult problem and in this case, the multi-parametric approach proved to be slightly better (88.9% accuracy) than the classification on individual feature spaces. The subsequent feature analysis confirmed that the key ROIs for patient identification do indeed belong to the temporal lobe. Their relevance in TLE has been indicated by clinical findings and similar research studies. These results reflect the sensitivity of quantitative imaging and the utility of the presented method towards the detection of TLE.

Classification models including PCA transformation after feature selection in general provided better results than the non-PCA models. Among the 8 classification models evaluated, the *anova-pca-linear* model demonstrated the best balance between classification performance and computation time, both key elements in the analysis large datasets. Although SVM behaves coherently in small sample scenarios, a larger patient sample would allow the amount of internal cross-validation and sub-sampling to be reduced. In machine learning terms, this could lead to a better balance between bias and variance.

Bibliography

- [1] N. K. Focke, M. Yogarajah, M. R. Symms, O. Gruber, W. Paulus, J. S. Duncan, Automated MR image classification in temporal lobe epilepsy, *NeuroImage* 59 (1) (2012) 356 – 362.
- [2] S. Keihaninejad, R. A. Heckemann, I. S. Gousias, J. V. Hajnal, J. S. Duncan, P. Aljabar, D. Rueckert, A. Hammers, Classification and lateralization of temporal lobe epilepsies with and without hippocampal atrophy based on whole-brain automatic MRI segmentation, *PLoS ONE* 7 (4) (2012) e33096.
- [3] S. C. Deoni, B. K. Rutt, T. M. Peters, Rapid combined T1 and T2 mapping using gradient recalled acquisition in the steady state, *Magnetic Resonance in Medicine* 49 (3) (2003) 515–526.
- [4] S. C. L. Deoni, T. M. Peters, B. K. Rutt, High-resolution T1 and T2 mapping of the brain in a clinically acceptable time with DESPOT1 and DESPOT2, *Magnetic Resonance in Medicine* 53 (1) (2005) 237–241.
- [5] S. C. Deoni, High-resolution T1 mapping of the brain at 3T with driven equilibrium single pulse observation of T1 with high-speed incorporation of RF field inhomogeneities (DESPOT1-HIFI), *Journal of Magnetic Resonance Imaging* 26 (4) (2007) 1106–1111.
- [6] S. C. Deoni, Transverse relaxation time (T2) mapping in the brain with off-resonance correction using phase-cycled steady-state free precession imaging, *Journal of Magnetic Resonance Imaging* 30 (2) (2009) 411–417.
- [7] M. Jenkinson, S. Smith, A global optimisation method for robust affine registration of brain images, *Medical Image Analysis* 5 (2) (2001) 143 – 156.
- [8] M. Beg, M. I. Miller, A. Trounev, L. Younes, Computing large deformation metric mappings via geodesic flows of diffeomorphisms, *International Journal of Computer Vision* 61 (2) (2005) 139–157.

- [9] H. Huang, C. Ceritoglu, X. Li, A. Qiu, M. I. Miller, P. C. van Zijl, S. Mori, Correction of B0 susceptibility induced distortion in diffusion-weighted images using large-deformation diffeomorphic metric mapping, *Magnetic Resonance Imaging* 26 (9) (2008) 1294 – 1302.
- [10] S. C. Deoni, B. K. Rutt, T. M. Peters, Synthetic T1-weighted brain image generation with incorporated coil intensity correction using DESPOT1, *Magnetic Resonance Imaging* 24 (9) (2006) 1241 – 1248.
- [11] R. Desikan, F. Segonne, B. Fischl, B. Quinn, B. Dickerson, D. Blacker, R. Buckner, A. Dale, R. Maguire, B. Hyman, M. Albert, R. Killiany, An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest, *NeuroImage* 31 (2).
- [12] C. Cortes, V. Vapnik, Support-vector networks, *Mach. Learn.* 20 (3) (1995) 273–297.
- [13] B. E. Boser, I. M. Guyon, V. N. Vapnik, A training algorithm for optimal margin classifiers, in: *Proceedings of the fifth annual workshop on Computational learning theory, COLT '92*, ACM, New York, NY, USA, 1992, pp. 144–152.
- [14] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine Learning in Python , *Journal of Machine Learning Research* 12 (2011) 2825–2830.
- [15] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, *J. Mach. Learn. Res.* 3 (2003) 1157–1182.
- [16] B. Mwangi, T. Tian, J. Soares, A review of feature reduction techniques in neuroimaging, *Neuroinformatics* (2013) 1–16.
- [17] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning*, 2nd Edition, Springer Series in Statistics, Springer New York Inc., New York, NY, USA, 2001.

- [18] M. A. Hall, Correlation-based Feature Selection for Machine Learning, Tech. rep., University of Waikato (1999).
- [19] C. Ding, H. Peng, Minimum redundancy feature selection from microarray gene expression data, *Journal of Bioinformatics and Computational Biology* 03 (02) (2005) 185–205.
- [20] N. Sanchez-Marono, A. Alonso-Betanzos, M. Tombilla-Sanromn, Filter methods for feature selection -a comparative study, in: H. Yin, P. Tino, E. Corchado, W. Byrne, X. Yao (Eds.), *Intelligent Data Engineering and Automated Learning - IDEAL 2007*, Vol. 4881 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2007, pp. 178–187.
- [21] D. D. Cox, R. L. Savoy, Functional magnetic resonance imaging (fMRI) “brain reading”: detecting and classifying distributed patterns of fMRI activity in human visual cortex, *NeuroImage* 19 (2) (2003) 261 – 270.
- [22] V. Michel, E. Eger, C. Keribin, B. Thirion, Multiclass sparse bayesian regression for fMRI-based prediction, *Journal of Biomedical Imaging* 2011 (2011) 2:1–2:13.
- [23] V. Michel, A. Gramfort, G. Varoquaux, E. Eger, C. Keribin, B. Thirion, A supervised clustering approach for fMRI-based inference of brain states, *Pattern Recognition* 45 (6) (2012) 2041 – 2049.
- [24] J. Mourão-Miranda, A. L. Bokde, C. Born, H. Hampel, M. Stetter, Classifying brain states and determining the discriminating activation patterns: Support vector machine on functional MRI data, *NeuroImage* 28 (4) (2005) 980 – 995.
- [25] J. E. Jackson, *A User’s Guide to Principal Components*, John Wiley & Sons, Inc., 1991.
- [26] A. Golugula, G. Lee, A. Madabhushi, Evaluating feature selection strategies for high dimensional, small sample size datasets, in: *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, 2011, pp. 949–952.

- [27] G. Lee, C. Rodriguez, A. Madabhushi, Investigating the efficacy of nonlinear dimensionality reduction schemes in classifying gene and protein expression studies, *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 5 (3) (2008) 368–384.
- [28] H. Li, Z. Xue, M. Dulay, A. Verma, S. Wong, C. Karmonik, R. Grossman, S. Wong, Distinguishing left or right temporal lobe epilepsy from controls using fractional anisotropy asymmetry analysis, in: H. Liao, P. Edwards, X. Pan, Y. Fan, G.-Z. Yang (Eds.), *Medical Imaging and Augmented Reality*, Vol. 6326 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2010, pp. 219–227.
- [29] D. W. Gross, Diffusion tensor imaging in temporal lobe epilepsy, *Epilepsia* 52 (2011) 32–34.
- [30] M. Ahmadi, D. Hagler, C. McDonald, E. Tecoma, V. Iragui, A. Dale, E. Halgren, Side matters: Diffusion tensor imaging tractography in left and right temporal lobe epilepsy, *American Journal of Neuroradiology* 30 (9) (2009) 1740–1747.
- [31] A. Kalousis, J. Prados, M. Hilario, Stability of feature selection algorithms, in: *Data Mining, Fifth IEEE International Conference on*, 2005, pp. 8 pp.–.
- [32] A. Kalousis, J. Prados, M. Hilario, Stability of feature selection algorithms: a study on high-dimensional spaces, *Knowledge and information systems* 12 (1) (2007) 95–116.
- [33] P. Matykiewicz, J. Pestian, Effect of small sample size on text categorization with support vector machines, in: *Proceedings of the 2012 Workshop on Biomedical Natural Language Processing*, Association for Computational Linguistics, 2012, pp. 193–201.
- [34] F. Al Sufiani, L. C. Ang, Neuropathology of temporal lobe epilepsy, *Epilepsy research and treatment* 2012.
- [35] A. R. Khan, M. Goubran, S. de Ribaupierre, R. R. Hammond, J. G. Burneo, A. G. Parrent, T. M. Peters, Quantitative relaxometry and diffusion MRI for lateralization in MTS and non-mts temporal lobe epilepsy, *Epilepsy Research* 108 (3) (2014) 506 – 516.

- [36] S. J. Raudys, A. K. Jain, Small sample size effects in statistical pattern recognition: Recommendations for practitioners, *IEEE Transactions on pattern analysis and machine intelligence* 13 (3) (1991) 252–264.
- [37] J. Hua, Z. Xiong, J. Lowey, E. Suh, E. R. Dougherty, Optimal number of features as a function of sample size for various classification rules, *Bioinformatics* 21 (8) (2005) 1509–1515.
- [38] N. Kemmotsu, H. M. Girard, B. C. Bernhardt, L. Bonilha, J. J. Lin, E. S. Tecoma, V. J. Iragui, D. J. Hagler, E. Halgren, C. R. McDonald, MRI analysis in temporal lobe epilepsy: cortical thinning and white matter disruptions are related to side of seizure onset, *Epilepsia* 52 (12) (2011) 2257–2266.
- [39] M. Liu, L. Concha, C. Lebel, C. Beaulieu, D. W. Gross, Mesial temporal sclerosis is linked with more widespread white matter changes in temporal lobe epilepsy, *NeuroImage: clinical* 1 (1) (2012) 99–105.

Chapter 4

Classification of temporal lobe epilepsy patients using voxel-based asymmetry

4.1 Introduction

Several research studies have shown how the concept of *brain asymmetry* can be leveraged in the analysis of neurological conditions. For example, Pepe et al. [1] presents a framework for the validation of VBM measures of asymmetry by evaluating *difference images* obtained by comparing gray matter maps with their left-right flipped versions. Furthermore, Fang et al. [2] combines regions of interest (ROIs) extracted from a cortical atlas with white matter tractography (structural + DTI) to create an *adjacency matrix of brain connectivity* from which features are extracted to automatically classify between patients and patients with left and right TLE. In addition, Zhang et al. [3] defines the *global connectivity asymmetry* and the *pairwise brain-region synchronization* as two measurements of asymmetry between left-right pairs of homologous regions (as defined per an anatomical atlas) to show altered functional connectivity in epileptic patients using resting-state functional magnetic resonance imaging (fMRI).

Similarly, the experiments performed in previous chapters evidenced that asymmetry features lead classification efforts. Briefly:

- asymmetry features enabled the correct lateralization of patients with a *laterality score*

(Chapter 2, Equation 2.3)

- MTS+ are more asymmetric than MTS- patients (Chapter 2, Figure 2.8)
- L-TLE and R-TLE groups present distinct asymmetry patterns (Chapter 3, Figure 3.7).

with these observations in mind, this chapter focuses solely on asymmetry in the temporal lobe, exploring the concept of voxel-based asymmetry (VBA) instead of focusing on a list of anatomical structures as in previous chapters. VBA is independent of the original MR imaging intensities, enabling the integration of one or more MR image types under a single analysis technique. This chapter shows the construction of VBA maps, and discuss how these maps can be used to perform group-level statistical analyses as well as individual-level predictions.

4.2 Methods and Materials

The method is divided in four steps: (i) MRI acquisition and data preprocessing, (ii) registration to a symmetric template, (iii) generation of asymmetry maps; and (iv) group-level statistical analysis, and individual-level automatic classification. Figure 4.1 shows how these steps are connected.

4.2.1 Participants

21 patients with refractory temporal lobe epilepsy undergoing pre-surgical evaluation and were enrolled into this study. All patients had unilateral TLE which was confirmed by video EEG monitoring and/or clinical MRI examination which included 3D FLAIR coronal, T2w axial, 3D T1w axial, T2w gradient echo coronal, FSE-IR coronal and diffusion axial (6 directions, $b=1000$). Clinical details are available in Table 4.1. A group of 33 healthy control subjects was recruited as well. The mean age was 38.4 ± 12.6 (range 18-59, 11 female) for the patient group, and 28.4 ± 9 (range 20-54, 13 female) for the control group.

All the subjects underwent the same imaging protocol for this study as described in the following section. This project was approved by the Office of Research Ethics of Western

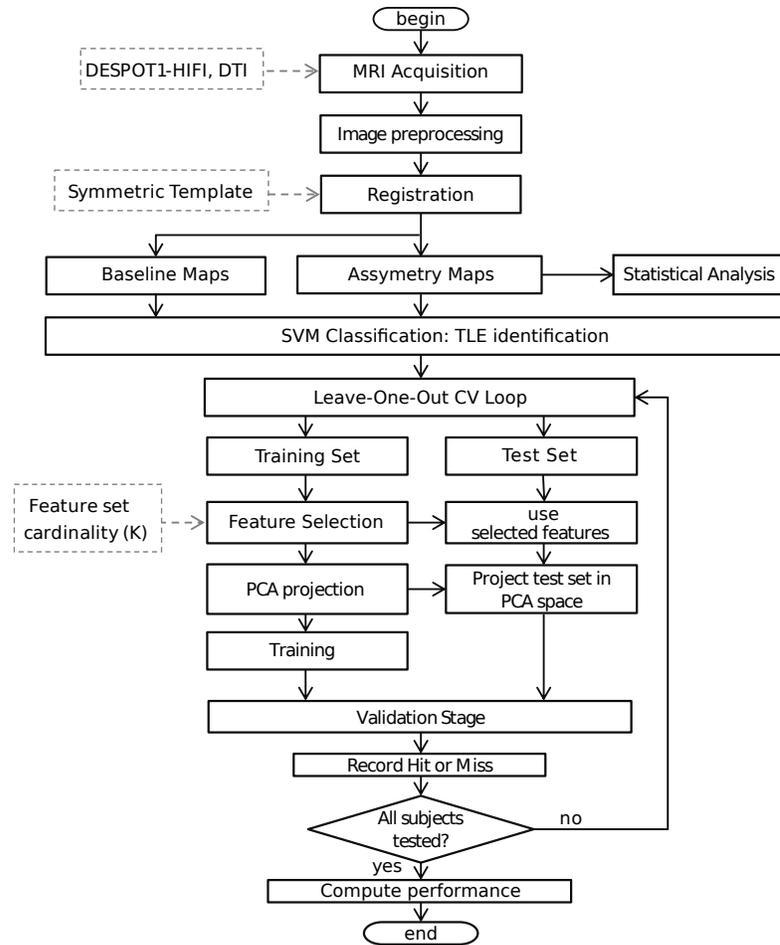


Figure 4.1: **Method overview.** Asymmetry maps are a personal signature that can be used for group-level (statistical analysis) and for individual-level prediction (automatic classification).

University (Canada) and all procedures were in accordance with the Helsinki Declaration of 1975, as revised in 2008.

4.2.2 Data acquisition and preprocessing

All subjects underwent imaging (relaxation mapping and diffusion-tensor imaging) on a 3T Discovery MR750 scanner (General Electric, Milwaukee, WI, USA) with a 32 channel head coil. We used the DESPOT1-HIFI method as described by Deoni et al. [4] to estimate T1 maps: Two SPGR images (flip angles: 4° and 18°) along with an additional inversion-prepared SPGR, all 1mm isotropic, were acquired. These images were registered to compensate for motion between the scans prior to the estimation of the T1 maps, with the FLIRT affine registration

Table 4.1: Clinical summary and demographics of the patients enrolled into this study.

ID	Age	Sex	Hand	Onset (years)	Duration (years)	Seizure Onset (EEG)	Anatomical MRI (1.5T)	surgery	Engel outcome
1	40	F	R	16	24	L	N	L-ATL	I
2	25	F	R	17	8	L	DNET	L-ATL	III
3	20	M	R	3	17	L	Left MTS	L-ATL	II
4	18	M	R	14	4	R	Right MTS	R-ATL	I
5	48	F	R	36	12	L	Left MTS	L-ATL	II
6	41	F	R	8	33	R	N	R-ATL	II
7	30	M	R	2	28	L	Left MTS	L-ATL	I
8	31	M	R	28	3	R	N	R-ATL	I
9	49	F	R	22	27	R	Right MTS, PVH, CD	R-ATL	I
10	32	F	R	19	13	L	Left MTS	L-ATL	I
11	56	F	R	15	41	R	N	R-ATL	II
12	43	F	R	3	40	R	Right MTS	R-ATL	I
13	23	M	R	18	5	L	N	L-ATL	I
14	34	M	L	15	19	L	Left MTS	L-ATL	III
15	33	F	R	20	13	R	Right MTS, PVLM	R-ATL	I
16	39	M	L	13	26	R	Right DNET or CD	R-ATL	I
17	59	M	R	11	48	L	Left MTS	-	-
18	58	F	R	8	50	L	possible right MTS	R-ATL	-
19	58	M	R	7	51	R	N	-	-
20	35	F	R	11	24	L	N	-	-
21	34	M	R	10	24	L	FCD or low-grade glioma	-	-

MTS: mesial temporal sclerosis; L:-left;R:-right; ATL: anterior temporal lobectomy; CD: cortical dysplasia;

DNET: dysembryoplastic neuroepithelial tumour; PVH: periventricular heterotopia; PVLM: periventricular leukomalacia;

N: normal MRI or non-specific findings.

tool [5] from the FSL library ¹, using the 18° image as reference. Diffusion tensor imaging (DTI) was performed using 41 gradient directions (b-value=1000) and 6 non-weighted (b=0) volumes (2.5 mm isotropic resolution). B0 susceptibility induced distortions were corrected using deformable registration of the average of non-weighted scans to the T1 map [6]. Eddy current correction and diffusion tensor estimation were performed with the FDT tool (part of the FSL library). The resulting maps of fractional anisotropy (FA) and mean diffusivity maps (MD). A synthetic bias-corrected T1-weighted image was created from the T1 map following the method described by Deoni et al. [7]. This T1-weighted image was required to perform the FA and MD registration to the T1 map. Skull stripping was performed using BET (also part of FSL). Brain masking and tissue segmentation were performed with FSL² [5], and registration to a *symmetric template* was performed with non-rigid B-spline registration, with affine initialization. Co-registered FA and MD maps were transformed to this symmetric template space, along with Jacobian-modulated gray matter density maps (GM*).

¹<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>

²FSL software: <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>

4.2.3 Measuring Voxel-based asymmetry

Under the hypothesis that focal morphological abnormalities can be identified by a comparison of *corresponding regions* between left and right hemispheres, four images per subject (T1, FA, MD, GM*) were analyzed. Each one of the images was flipped (left-right), followed by registration to the symmetric template. This step enabled a voxel-scale comparison between left and right hemispheres: the coordinates for any voxel A in the unflipped image are the same as the coordinates of its hemispherical opposite B in the flipped image (Figure 4.2).

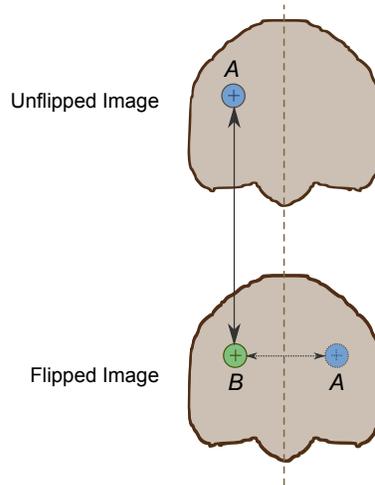


Figure 4.2: **Establishing hemispheric correspondence.** B is the corresponding region for A in the opposite hemisphere.

Defining homologous voxel neighbourhoods

It is unrealistic to obtain a perfect voxel-by-voxel correspondence between left and right hemispheres: anatomical variability due to differences in cortical folding and normal occurring structural changes need to be considered. Similarly, the registration to the common space provided by the symmetric template has its own limitations: registration may not be perfect due to anatomical variability, partial volume effects and signal to noise ratio. Also, the voxel size before registration is not the same between relaxation mapping (1mm) and diffusion tensor images (2.5mm). To account for these constraints, the comparison of homologous left-right regions was formulated between spherical voxel neighbourhoods with a parametric radius r in

voxels (in addition to the central voxel). Neighbourhoods of different sizes were studied: small neighbourhoods ($r = 1, r = 2$ voxels) provided very little data to produce a good comparison, whereas large neighbourhoods permitted a better estimation of left-right differences, and compensated better for the aforementioned constraints. Nonetheless, large neighbourhoods ($r \geq 4$ voxels) were also computationally expensive and reduced the localization ability of small differences. Under these considerations, a radius of $r = 3$ voxels was selected, encompassing a volume of 123mm^3 . This equated to 123 voxels for the T1 and GM* images, and 7.8 voxels for the FA and MD images, before registration.

Measuring voxel-based asymmetry

Focal left-right differences can be quantified by introducing a *voxel-based asymmetry* estimation [L]. For a given neighbourhood defined by voxel v (and radius $r = 3$) two corresponding regions are extracted from the unflipped (U) and flipped (F) images, namely regions $U(v)$ and $F(v)$.

These regions are compared using a two-sample Kolmogorov-Smirnov (KS) test. Thus, voxel-based asymmetry is defined as the result of this test:

$$L_{s,i}(v) = KS(U_{s,i}(v), F_{s,i}(v)) \quad (4.1)$$

where s is a subject and i is one of four possible image sources: T1, FA, MD, GM*.

The KS test measures the empirical cumulative distribution functions (ECDFs) of both regions (U,F) and returns the maximum difference between them (Figure 4.3). Since ECDFs are bounded in the $[0,1]$ interval the minimum difference returned is zero when there is a complete overlap (100% symmetry) and the maximum is one when there is no overlap (100% asymmetry).

Subsequently, *voxel-based asymmetry maps* [A] are generated by computing voxel-based asymmetry for across all possible voxel neighbourhoods (background voxels are not evaluated). Formally:

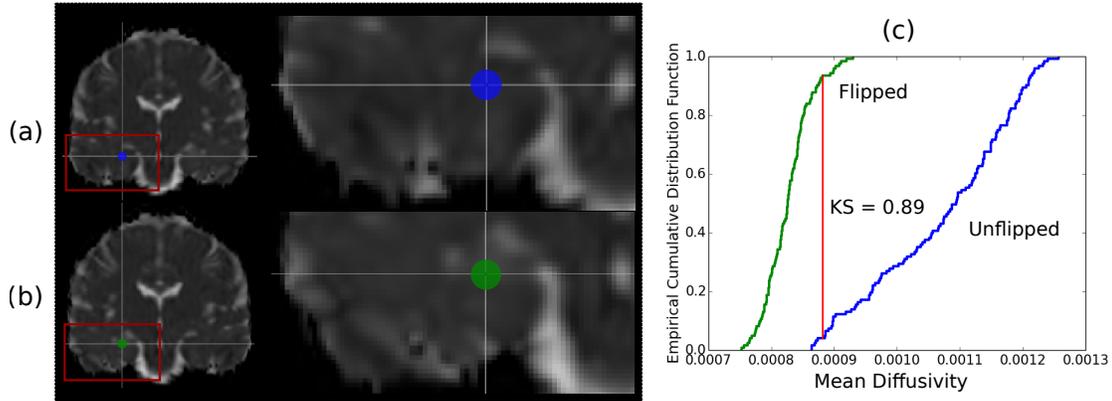


Figure 4.3: **Evaluation of voxel-based asymmetry.** (a) Original image; (b) flipped and co-registered image; (c) A two-sample Kolmogorov-Smirnov (KS) test is performed between homologous (left-right) neighbourhoods. In this example, the KS test is performed on a mean diffusivity (MD) image in a neighbourhood closer to the hippocampus for a patient diagnosed with mesial temporal sclerosis. The red line indicates the outcome of the test which corresponds to the maximum distance between the left and right empirical cumulative distributions, indicating that the signal in these regions are highly asymmetric.

$$A_{s,i} = \{L_{s,i}(v) : \forall v \in i\} \quad (4.2)$$

Four asymmetry maps were obtained per participant as depicted by Figure 4.4.

A visual inspection (Figure 4.5) of the maps using a threshold of 70% ($L = 0.7$) revealed two things: (i) each subject has a particular *asymmetry signature*, and (ii) TLE patients seemed to have more asymmetry in the temporal lobe.

A statistical analysis was carried out to test this hypothesis, dividing the temporal lobe in three regions (Figure 4.6):

- MSL: The mesial temporal region including the hippocampus, amygdala and adjacent mesial structures.
- NEO: A region that encompasses the temporal neocortex, cortex and white matter.
- ATL: The region constructed by the combination of MSL and NEO, corresponding to the temporal lobe.

These regions are selected because TLE can be either mesial or lateral. The statistical

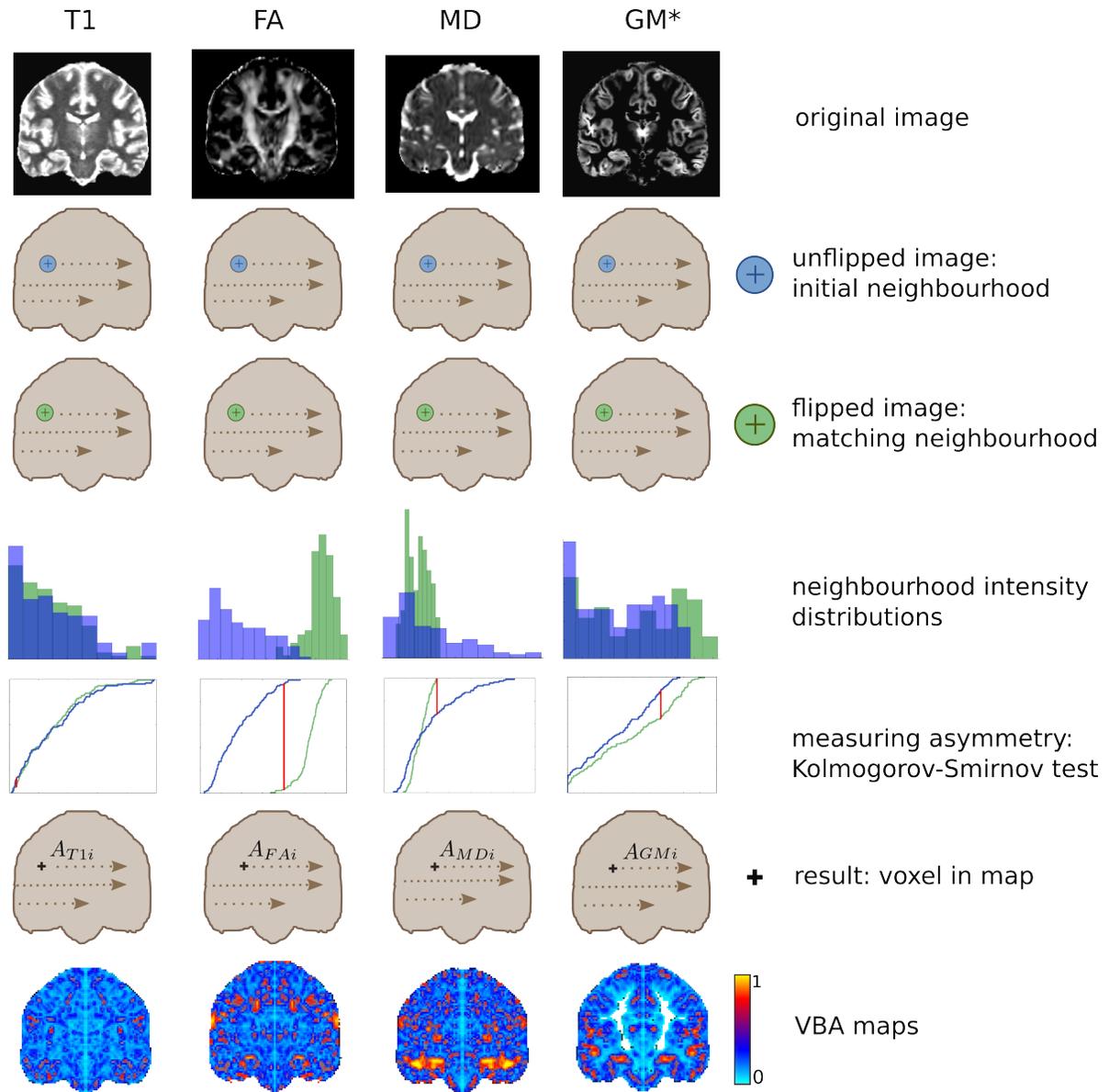


Figure 4.4: **Construction of voxel-based asymmetry maps (VBA maps)**. A pair of images (flipped, unflipped) are obtained by registration using the symmetric template. Voxel neighbourhoods with the same coordinates in these images represent opposite regions between cerebral hemispheres. For each pair of regions the intensity distribution is collected and a measure of *voxel-based asymmetry* is obtained by performing a two-sample Kolmogorov-Smirnov test. The result of the KS test is stored in the neighbourhood central voxel. Repeating this procedure for all brain voxels generates VBA maps.

analysis in these regions can establish the type of asymmetry that is most significant to identify patients from controls. Additionally, since our patient cohort includes patients with MTS, it is interesting to see if there could significant differences in asymmetry outside the mesial region.

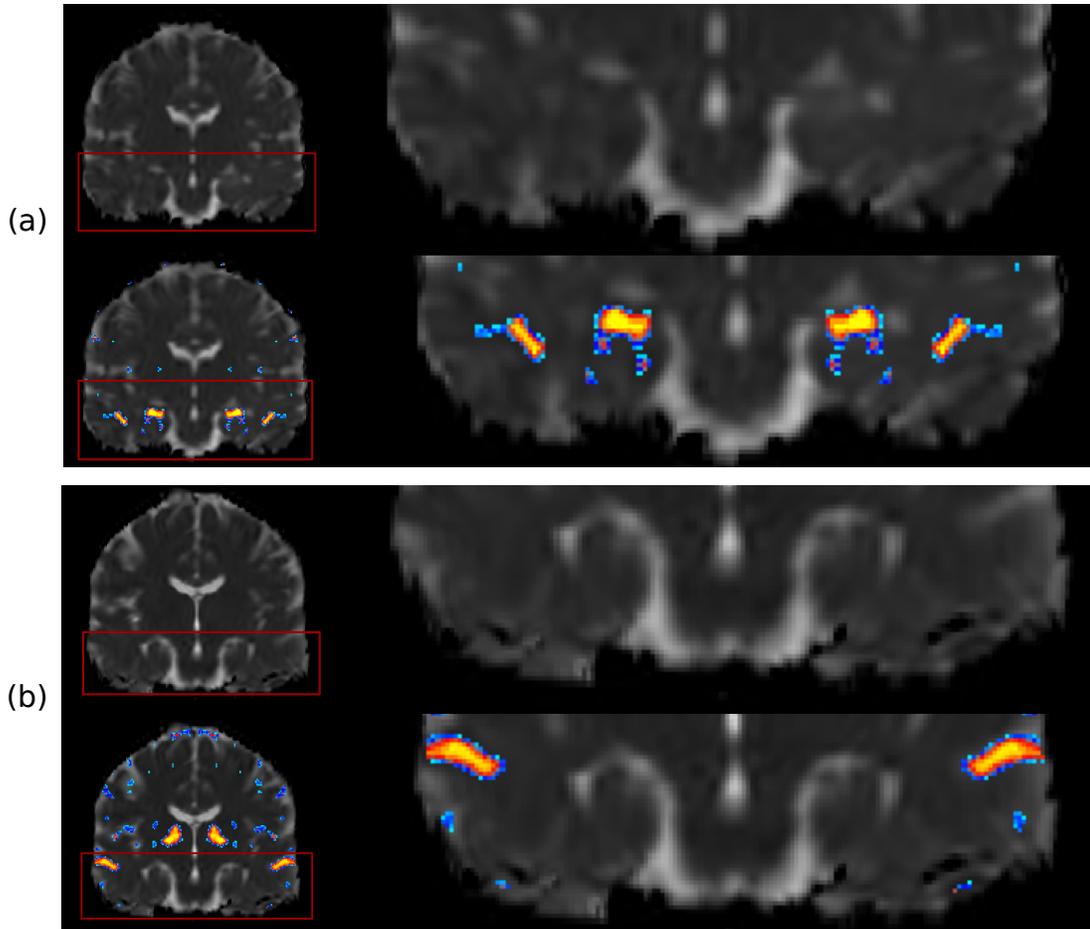


Figure 4.5: **Visual inspection of asymmetry maps.** (a) TLE patient with mesial temporal sclerosis; (b) healthy volunteer; both maps have been thresholded to show voxel-based asymmetries in the range of 70% (blue) to 100% (yellow) only. The red box shows the region that has been zoomed in on the right. It was determined that each subject (patients and controls) had a particular *asymmetry signature* by examining these asymmetry maps. This signature *by itself* was not an indication of a pathological process. It was only through comparisons at the group or individual-to-group levels that *abnormal asymmetry* were established. For instance, using a MANOVA design, we corroborated the observation that TLE patients in our sample have abnormally high (multivariate) asymmetry in the temporal lobe.

4.2.4 Statistical analysis of temporal lobe asymmetry

With the purpose of performing group statistical analyses, *temporal lobe asymmetry* (δ) was defined as the average of voxel-based asymmetry in the temporal lobe. Formally:

$$\delta_{s,i}(R) = \frac{\sum_{v \in R} L_{s,i}(v)}{|R|} \quad (4.3)$$

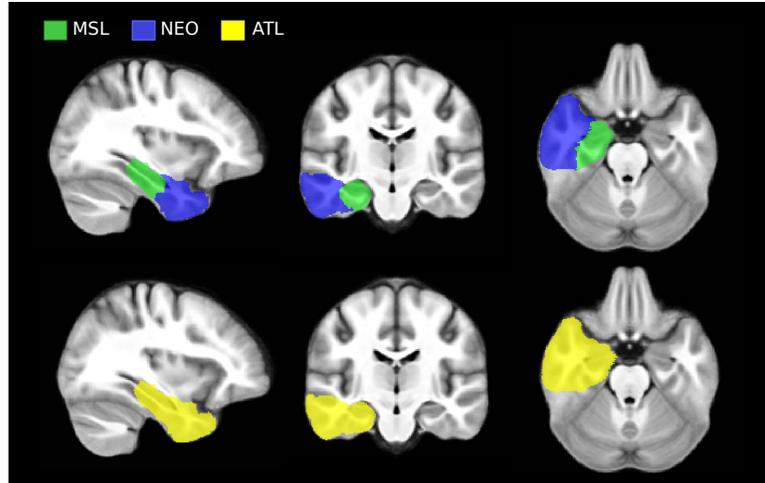


Figure 4.6: **Temporal lobe regions.** The information contained in the asymmetry maps is mirrored between the left and right hemispheres. Therefore, only one hemisphere is required for feature extraction. That is not the case for the baseline experiment (discussed in 4.2.6). MSL: mesial; NEO: lateral; ATL: anterior temporal lobe.

where R is one of three possible temporal regions: MSL, NEO, or ATL; and $|R|$ is the number of voxels in region R . A multivariate statistical analysis was performed using SPSS³ to test the hypothesis that there would be one or more differences in temporal asymmetry means between patients and controls.

4.2.5 TLE patient identification using Support Vector Machines

To determine the applicability of the voxel-based asymmetry metric on the detection of TLE, a linear support vector machine (SVM) was employed. A linear SVM is a supervised binary classifier that computes the hyperplane or *decision boundary* that maximizes the separation (margin) between two groups [8]. The SVM was trained/tested using leave-one-out cross-validation (LOOCV) due to the small sample size. The SVM analysis was done using the *scikit-learn* library [9].

One classification experiment was performed for each region of interest (Figure 4.6). In each experiment, the classifier was restricted to that region for extraction and selection of features. This allowed us to establish and compare the sensitivity and specificity of the proposed

³IBM Corp. Released 2013. IBM SPSS Statistics, Version 22. Armonk, NY: IBM Corp.

method for TLE patient identification using features belonging to the mesial temporal region (MSL), the lateral temporal region (NEO), and the anterior temporal lobe (ATL).

Feature selection

Feature vectors were constructed for each subject by concatenating the asymmetry maps in the temporal lobe regions as follows:

$$\mathbf{V}_{s,\varphi_R} = [\varphi_R(A_{s,TI}), \varphi_R(A_{s,FA}), \varphi_R(A_{s,MD}), \varphi_R(A_{s,GM*})] \quad (4.4)$$

where $\varphi_R(x)$ is a function that returns the voxels in R , with $R \in \{MSL, NEO, ATL\}$ in a VBA map x . Since all the VBA maps are registered to the symmetric template, the respective regions R have the same number of voxels, and thus, the feature vectors have the same length.

Binary class labels were assigned to each feature vector \mathbf{V} (0 = patient, 1 = control). Then, the classifier was trained and tested using leave-one-out cross-validation (LOOCV) after which performance metrics such as accuracy, sensitivity and specificity were collected.

Each training set, as generated per LOOCV, was randomly sampled 5 times using stratified shuffling and splitting to generate *feature selection folds*. Each fold retained 90% of the subjects in the training set. In addition, the stratified shuffling and splitting enforces the ratio between patients and controls to be similar to that of the training set being sampled. This creates balanced folds and strengthens the search for discriminative features.

An ANOVA-based feature selection algorithm was executed on each fold and the top K most discriminative features were retained. Feature ranking was based on the resulting F-value and ties were broken randomly. This procedure created a feature set candidate for each fold. Then, the final feature set was assembled by features that were consistently selected across all feature folds (Figure 4.7).

Once features were selected, the training dataset was transformed using principal component analysis (PCA) and the 5 top components were kept, obtaining a compact data representation. The same features were selected in the testing set, which was posteriorly projected onto PCA space (Figure 4.1).

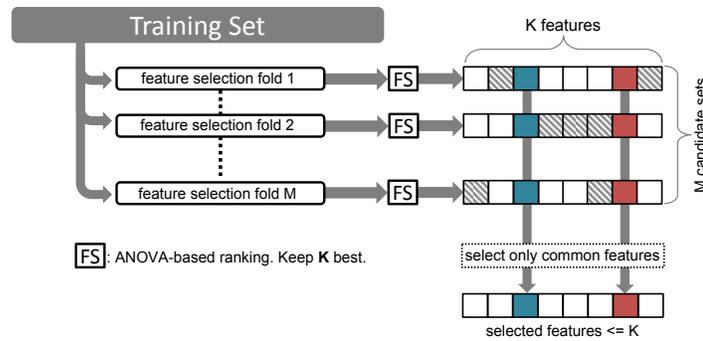


Figure 4.7: **Feature selection.** The algorithm randomly samples the training set, creating feature selection folds ($M=5$). Then, features on each fold are ranked using an ANOVA test and only the K top features are kept. Finally, only those features identified in all selection folds are chosen to constitute the definitive feature set that will be used in training and testing

Robust features and informative features

A feature set is obtained on each iteration of the LOOCV loop. This feature set can be different from iteration to iteration due to the fact that the training set is different every time and also due to the random sampling mechanism. Therefore it is important to determine which features are recurrent over all classification iterations. These *robust features* are obtained by examining the feature sets obtained on all the iterations of the LOOCV loop and retaining only those that consistently appear in all of them. This mechanism is similar to that described for the construction of the final feature set on every iteration, however in this case, the selection of features occurs across all LOOCV feature sets instead of feature folds in a particular training set.

Given that the feature selection algorithm requires setting the K parameter (cardinality) *a priori*, the performance of the classifier may vary depending on how large the set of features used for classification is. Therefore, the identification of robust features is limited by the selection of K . To account for this free parameter, robust features were evaluated for a range of K values, starting at 1000 features, incrementing by 1000 features every time, until reaching the maximum number of available features T . Then, *informative features* were identified as those features that were consistently selected at least 80% of the times across all cases where the classifier performed well (above the measured average accuracy). An analysis of informative

features is presented in the results section.

4.2.6 Baseline experiment

A baseline experiment was proposed to compare the effectiveness of TLE identification based on the idea of focal asymmetry features versus classification based on the information contained in the T1, FA, MD and GM* images directly (no asymmetry calculation). The feature extraction and feature selection steps for this experiment were similarly performed on the MSL, NEO and ATL regions. However in this case, both the left and right hemisphere are explored for features. A linear SVM, identical to the one used in the VBA classification, was trained and tested following the same methodology as described before.

4.3 Results

4.3.1 Differences in temporal lobe asymmetry between patients and controls

The descriptive statistics associated with temporal lobe asymmetry between patients and controls are reported in Table 4.2. It can be seen that the temporal lobe asymmetry mean in the TLE group is higher than in the control group, except by the gray matter density (GM*) in the NEO and ATL regions.

Table 4.2: **Descriptive statistics.** Average asymmetry and standard deviation per per group and region of interest

ROI	Group	T1		FA		MD		GM*	
		μ	σ	μ	σ	μ	σ	μ	σ
MSL	Control	0.22	0.02	0.31	0.03	0.33	0.04	0.31	0.03
	TLE	0.23	0.02	0.33	0.03	0.39	0.07	0.33	0.04
NEO	Control	0.26	0.02	0.36	0.02	0.37	0.03	0.33	0.03
	TLE	0.29	0.03	0.38	0.03	0.43	0.06	0.33	0.03
ATL	Control	0.25	0.02	0.33	0.02	0.36	0.03	0.33	0.02
	TLE	0.27	0.02	0.37	0.03	0.42	0.06	0.33	0.03

A one-way multivariate analysis of variance (MANOVA) was conducted, to test the hypothesis that there would be mean differences between patients and controls in terms of temporal

lobe asymmetry. For this test, the subject group (patient,control) was considered as the independent variable, and the 4 regional asymmetries, δ_{T1} , δ_{FA} , δ_{MD} , δ_{GM^*} (Equation 4.3) were then identified as dependent variables. Statistical significant differences were identified in the mesial, lateral and anterior temporal regions (Table 4.3). This suggested that at least one type of temporal lobe asymmetry was different between patients and controls on these regions.

Table 4.3: **Multivariate tests results.** ROI asymmetry between TLE patients and control is significant in the regions of interest: MSL($p < 0.05$), NEO($p < 0.001$), and ATL($p < 0.001$); BT: Box's test of equality of covariance matrices; Λ is Wilks' Lambda.

ROI	BT	Λ	F	df_1	df_2	Sig.	η^2
MSL	0.054	0.793	3.189	4	49	0.021	0.207
NEO	0.001	0.594	8.359	4	49	0.000	0.406
ATL	0.013	0.606	7.967	4	49	0.000	0.394

To identify regional asymmetries contributing to the MANOVA results, a follow-up ANOVA test was performed on each dependent variable. For these tests, the significance level α of 0.5 selected and posteriorly adjusted by the number of ANOVA tests (Bonferroni correction). Thus the significance level for each individual ANOVA was set at $\alpha/n = 0.0125$. The results of these follow-up tests indicated that in the mesial region only the MD asymmetry was significantly different between patients and controls ($p = 0.001$); whereas in the lateral and ATL regions, T1, FA and MD asymmetries were significantly different ($p < 0.001$); GM* asymmetry was not significant in any of the three regions (Figure 4.8).

4.3.2 TLE patient identification using SVM on voxel-based asymmetry features

Identification of TLE patients with the VBA classifier was robust, obtaining accuracies between 87% and 92% (Table 4.4). The overall best performance was obtained by the VBA classifier in the lateral region (accuracy=91%, sensitivity=86%, specificity=94%). Figure 4.9 shows a three-dimensional representation of the feature space with the respective SVM decision boundary showing four distinct scenarios. Each one of these cases correspond to an different iterations in the LOOCV loop, where the feature space is slightly different each time

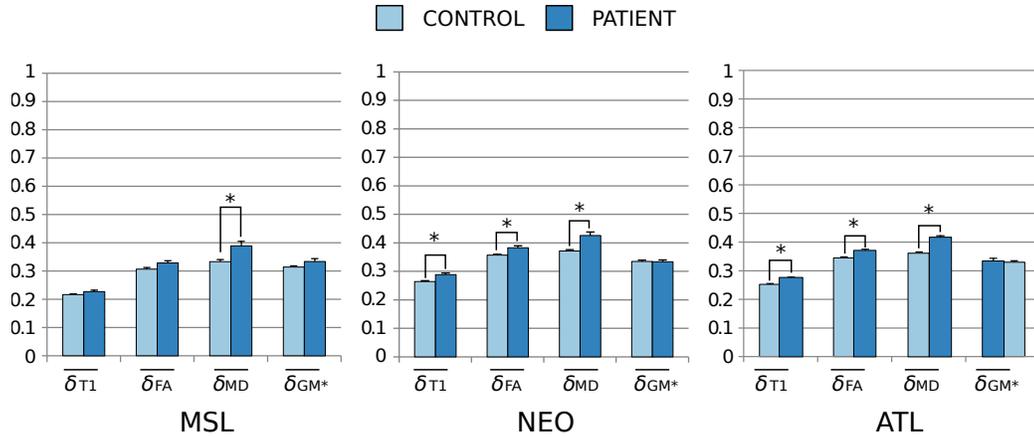


Figure 4.8: **ROI asymmetry differences between TLE patients and controls.** These results indicate that temporal lobe asymmetry in TLE patients is higher than in healthy volunteers for T1, FA and MD in the lateral and ATL regions, and only on MD for the mesial region. These results were obtained using a set of follow-up ANOVAS (after performing a MANOVA test per region that considered differences between TLE patients and controls in all four temporal lobe asymmetries jointly). For these ANOVA tests the statistical significance was set at $p < 0.0125$, which accounted for multiple comparisons.

and so it is the decision boundary. This low-dimensional projection shows that the control group tends to form a tight cluster whereas the patient group is much more scattered in terms of *multivariate voxel-based asymmetry*.

The VBA classifier outperformed the baseline classifier. Each experiment was repeated, increasing the value of the requested feature set cardinality K , up to the total number of available features T (size of the respective ROI in voxels). Table 4.4 summarizes the results for the best and average performances of both experiments (Baseline, VBA).

Figure 4.10 shows the accuracy of the classifier in function of the feature set cardinality K . In all three ROIs, the VBA classifier obtained an optimal performance with considerably less number of features. This indicates that voxel-based asymmetry is more informative than raw voxel intensities for the task of patient identification. Average performances are indicated dotted lines showing that the VBA classifier *generally* outperformed the baseline classifier.

4.3.3 Analysis of informative features

Most informative features across all feature set cardinalities were identified as follows. For a feature f and the robust feature set F_K determined by K , the indicator function $\mathbf{1}(f, K)$ is

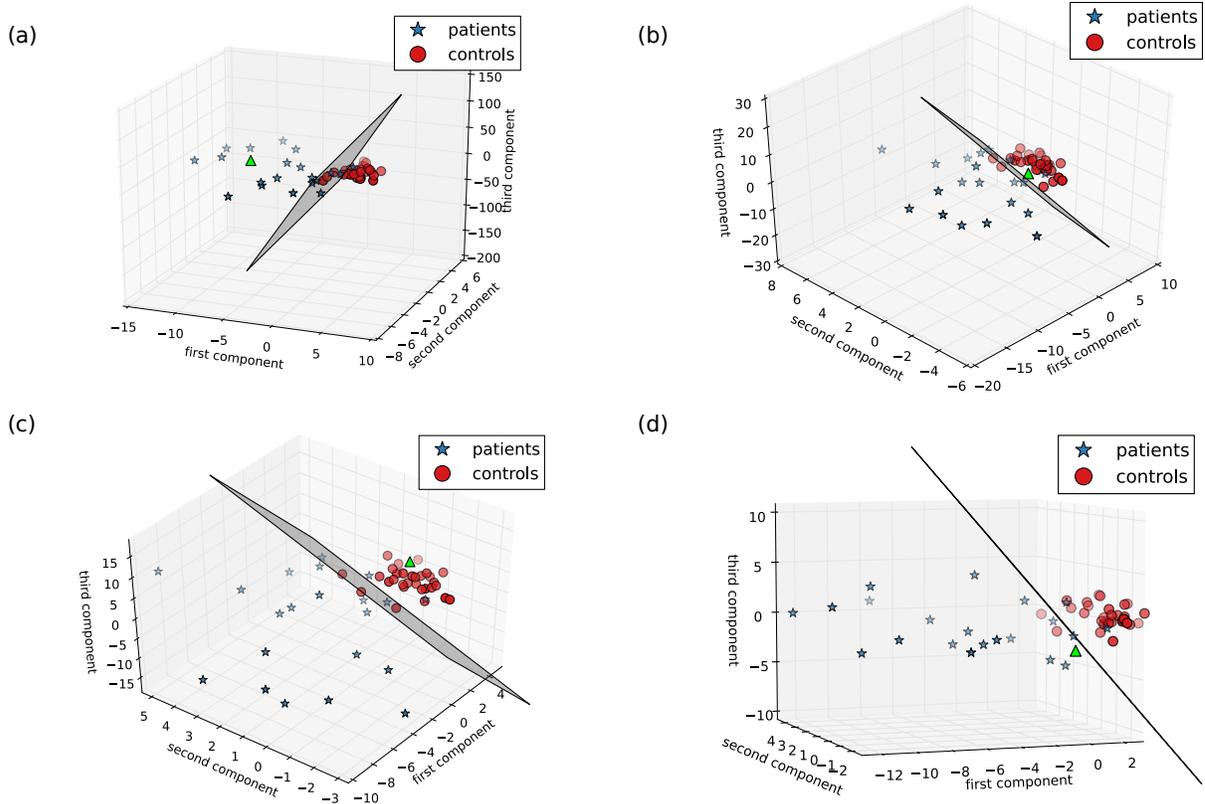


Figure 4.9: **SVM classification.** This figure shows a PCA projection of the high-dimensional feature space (asymmetry features), with the respective decision boundary determined by the SVM, in several scenarios: (a) a patient being correctly classified; (b) a patient being misclassified; (c) a control being correctly classified; and (d) a control being misclassified. Each case represents an iteration of the LOOCV loop, where the subject being classified (green triangle) is excluded from training. In all cases, the control group presents a cluster tendency whereas patients are sparse in terms of multivariate asymmetry.

defined as:

$$\mathbf{1}(f, K) = \begin{cases} 1 & : f \in F_K \\ 0 & : f \notin F_K \end{cases} \quad (4.5)$$

Then a frequency function $Q(f)$ can be defined as:

$$Q(f) = \frac{\sum_{K=1000}^T \mathbf{1}(f, K)}{N} \quad (4.6)$$

where N is the number of times the SVM performed with the measured average accuracy or higher. Then, f was determined to be *informative* if its frequency $Q(f)$ was higher than a set

Table 4.4: **Identification of TLE patients using SVMs.** Automatic classification using voxel-based asymmetry maps (VBA) outperformed the baseline experiment, which employed voxel-intensity features from the original images (Base).

	MSL		NEO		ATL	
	Base	VBA	Base	VBA	Base	VBA
Best Performance						
Accuracy	83%	87%	77%	92%	81%	87%
Sensitivity	67%	81%	71%	86%	71%	71%
Specificity	94%	91%	81%	94%	88%	97%
Average Performance						
Accuracy	73%	79%	60%	78%	67%	80%
Sensitivity	63%	73%	49%	79%	54%	72%
Specificity	79%	83%	66%	77%	76%	86%

threshold of 0.8. In other words, when it was selected at least 80% of the times across all good classification cases.

Table 4.5 shows the results of this analysis. The coverage metric C gives the proportion of informative features to the total number of available features T , which corresponds to the size of the respective ROI. Since this metric takes into account only those cases where the classifier performed at or above average, it shows that in general very little information is required to do so. However, the coverage metric must be analyzed in the context of how good the classifier performed (average accuracy included in the table): notably 24% of mesial features are informative while only 2% and 9% in the lateral and ATL regions are informative for classification performances around 80% for the VBA classifier. This could indicate that voxel-based asymmetry in the mesial temporal lobe is more widespread than in the peripheral white matter, gray matter and neocortex for TLE patients. In other words, more informative asymmetry features are found in the mesial region than outside of it in average for all cases where the classifier performed with an accuracy of 80% or above.

A composition of informative features is also provided in Table 4.5 in terms of their origin. Interestingly, MD features are the most common among informative features for the VBA classifier. In the mesial region they account for 35% of the total of informative features, while their percentages are 48% and 43% in the lateral and ATL regions respectively. A similar

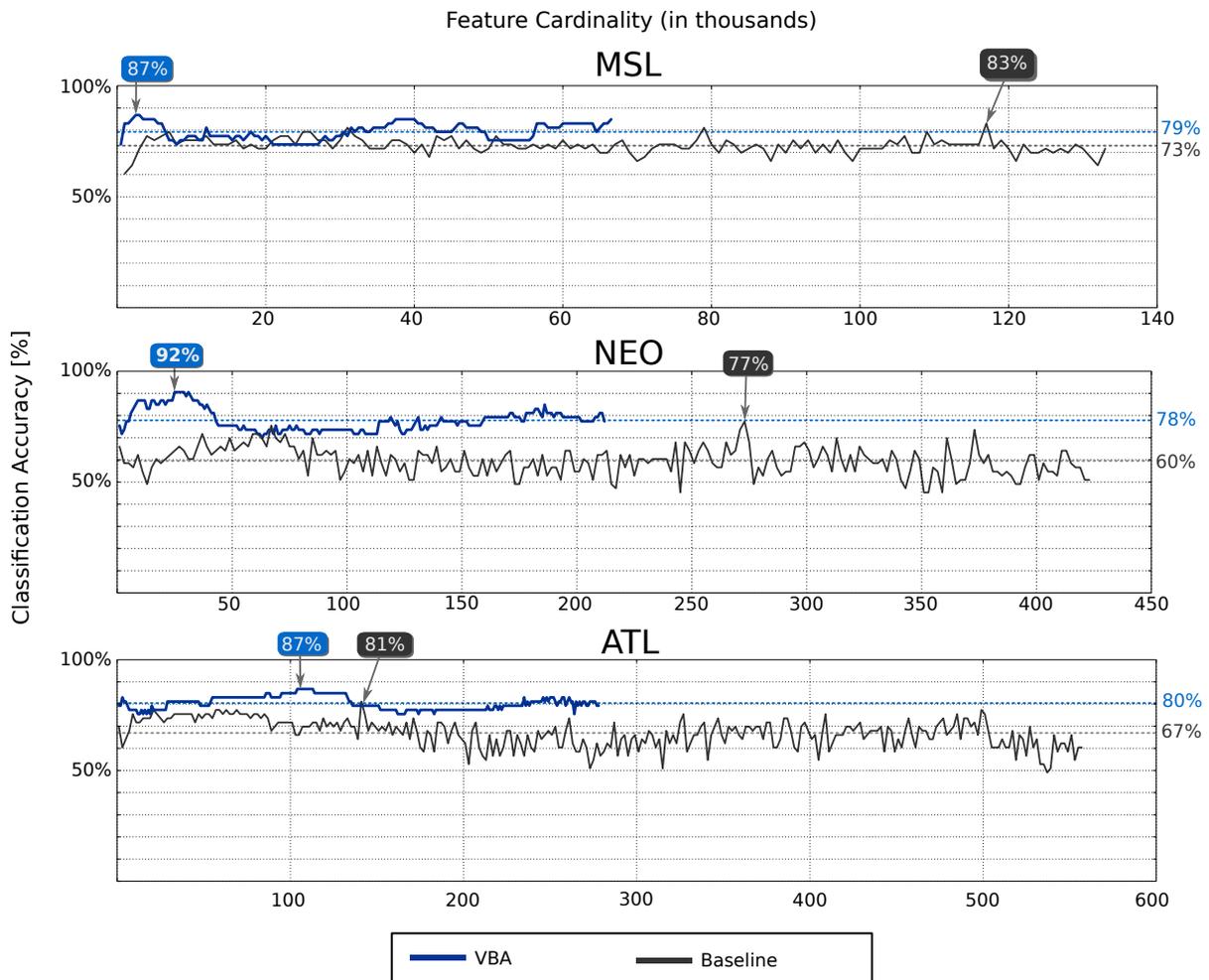


Figure 4.10: **Classification accuracy as a function of feature set cardinality.** In all three ROI, the VBA classifier outperformed the baseline classifier. These plots also show that classification accuracy for the VBA classifier is relatively more stable than for the baseline classifier.

result was observed when analyzing the composition of the robust feature set associated to the best classification cases of the VBA (not shown in the table): 75% of the features in the MSL region were MD features, 47% in NEO and 33% in the ATL region. These results point to the relevance that mean diffusivity images and their potential in the identification of patients with TLE.

Informative features in the mesial temporal region (MSL)

Informative features for the identification of patients with TLE were found in the hippocampal formation (Figure 4.11a). MD asymmetry in the hippocampus was the predominant feature

Table 4.5: **Analysis of informative features.** Informative features are consistently selected at least 80% of the times (see 4.2.5)

	MSL		NEO		ATL	
	Base	VBA	Base	VBA	Base	VBA
informative features [M]	4,725	16,026	13,124	4,800	13,107	25,784
available features [T]	133,224	66,612	424,024	212,012	557,284	278,624
coverage [C] = [M]/[T]	4%	24%	3%	2%	2%	9%
average SVM accuracy	73%	79%	60%	78%	67%	80%
Feature origin						
T1	12%	22%	11%	26%	11%	24%
FA	23%	17%	18%	21%	20%	20%
MD	7%	35%	6%	48%	7%	43%
GM*	6%	27%	22%	5%	18%	13%

among all four kinds of asymmetry (T1, FA, MD, GM*). This finding is consistent with the ROI asymmetry analysis where MD asymmetry showed to be different with statistical significance between patients with TLE and healthy control volunteers. MD and FA features extend anteriorly including parts of the amygdala, the hippocampus, the entorhinal cortex, the fusiform gyrus, and temporal pole.

Correspondingly, T1 and GM* informative features follow the CA1 hippocampal region towards the subiculum, while anteriorly, the T1 and GM* changes extend towards the entorhinal cortex and the temporal pole. These findings could be indicative of common areas of cortical integrity disruption in TLE patients.

Informative features in the lateral temporal region (NEO)

In contrast to mesial findings, common voxel-based asymmetry patterns in the lateral temporal region appear primarily in the white matter. Figure 4.11b shows changes along the temporooccipital inferior longitudinal fasciculus and frontotemporal uncinate and arcuate fasciculi connections with a predominance of MD findings followed by T1 changes.

On the other hand, FA and GM* changes are minor, appearing scattered in the superior, medial and inferior gyri.

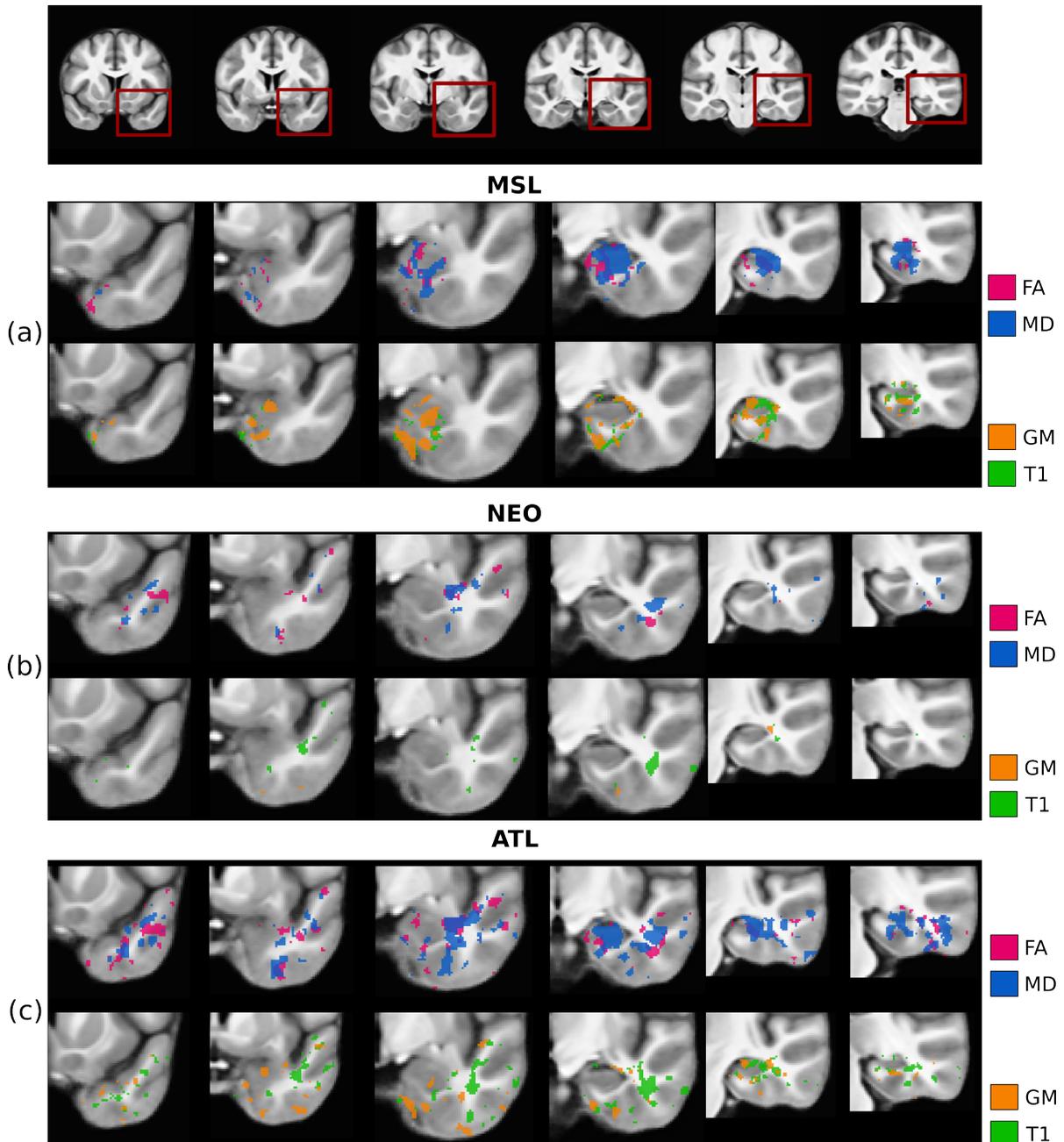


Figure 4.11: **Analysis of informative temporal lobe features.** (a)MSL; (b)NEO; (c)ATL. GM* is gray matter density (not volume). The features are overlaid on the symmetric template.

Informative features in the anterior temporal lobe (ATL)

When the method is not restricted to either the mesial or the lateral regions, but instead is allowed to select features from the entire anterior temporal lobe (ATL), the distribution of features resembles the patterns observed on the individual MSL and NEO regions (Figure 4.11c):

MD asymmetry is salient in the hippocampus, parahippocampal gyrus and temporal white matter, with some minor clusters forming in the inferior and medial temporal cortices. FA changes are scattered, while gray matter density changes occur primarily in the parahippocampal, fusiform and inferior gyri and project anteriorly onto the temporal pole. T1 asymmetry is present in the hippocampus and the temporal white matter with some isolated small clusters in the temporal cortex.

4.4 Discussion

This study presented a new approach for the detection of TLE. We proposed a metric of voxel-based asymmetry varying between zero (total symmetry) and one (total asymmetry) thus, being independent of the original unit of measurement, reporting complete symmetry as zero and complete asymmetry as unity in a continuous range. This metric enabled the integrative analysis of dissimilar magnetic resonance techniques, specifically, of quantitative relaxometry (T1), diffusion tensor imaging (MD, FA) and voxel-based morphometry (GM*).

Two different kinds of analysis were performed using the resulting asymmetry maps. Firstly, a multivariate analysis of variance identified relevant types of temporal lobe asymmetry that were significantly different between TLE patients and healthy volunteers, pointing to MD asymmetry as the most significant difference in the mesial region; T1, FA and MD changes were also evidenced in the lateral and ATL regions; and no significant difference was found in terms of GM* asymmetry. Secondly, automatic classification using a support vector machine (SVM) was performed using voxel-based asymmetry maps. The SVM performance was then compared to the performance obtained by using the original images (T1, FA, MD, GM*) with no computation of voxel-based asymmetry (baseline experiment). The results conveyed a superior performance in terms of classification accuracy, sensitivity and specificity for the SVM trained on voxel-based asymmetry maps.

4.4.1 Comparison to multi-voxel pattern analysis methods

Multivariate voxel pattern analysis (MVPA) describes a set of analysis techniques that explore voxel-based voxel-neighbourhood patterns in four stages: (i) feature selection; (ii) pattern assembly; (iii) classifier training and, (iv) generalization testing [10]. The unit of analysis in these studies is the *searchlight* [11], which is a spherical voxel-neighbourhood, similar to the one used in our method for the calculation of asymmetry maps. In contrast to MVPA, the training and generalization test steps in our method occur at the brain-level or in selected ROI: the spherical neighbourhoods are used to generate a *global asymmetry map* that is then used for classification. In MVPA, (iii) and (iv) occur at the voxel-neighbourhood level. The performance of the classifier performance is used then as an indication of the discrimination between two experimental conditions *in that region* [12]. Another important distinction is that of purpose: voxel-based asymmetry creates individual signatures of left-right hemispherical differences to study abnormal asymmetric conditions such as the pathologies associated to TLE; whereas MVPA is general purpose, and has been commonly used in the study of fMRI [13].

4.4.2 Comparison to similar multi-parametric classification studies

Our results were comparable to those obtained by Zhang et al. [3] which performed automatic TLE patient detection using rest-state fMRI features (global connectivity asymmetry, pairwise brain-region synchronization). Their reported average accuracy range was 75.5%-75.8%, while the average accuracy of our method was 78%-80%. Similarly, their best accuracy was 83.9%, whereas our results were in the 87%-92% range.

Analogously, Focke et al. [14] performed multi-parametric automatic classification of TLE patients obtaining accuracies in the 90%-97% range, with three main differences: (i) classification was performed individually on gray matter, white matter, T2, FA and MD maps; (ii) their patient sample was larger and more homogeneous (38 patients with hippocampal sclerosis and no other lesions); (iii) classification was optimized by weighing individual images in the training set with group-level differences (SPM F-contrasts). Non-optimized results are

also reported, with accuracies going down to 69%, which interestingly resembles the behaviour evidenced in our baseline experiment. Additionally, their best classification scenarios were obtained by training the SVM on MD features. This finding echoes the MD relevance observed by discriminating the origin of the features in our integrative design.

4.4.3 Feature selection trade-offs

There are two alternatives when it comes to small datasets subjected to a LOOCV classification scheme: to not perform feature selection, or perform feature selection *once for every training set*. An argument for the first approach is that applying feature selection in a small dataset could pick up spurious relationships in the data rather than “good” features. This is possible if real correlations in the sample are weaker than the spurious ones. The no feature selection approach is followed by Focke et al. [14]. Though they applied a mechanism that effectively weighs down irrelevant features, all the features are used in training and testing of their classifier. In contrast, if there is no feature selection at all, there is no certainty that the classifier is identifying a real effect instead of a sample effect. The lack of feature selection could impact the classifier’s generalization ability if the sample under consideration is small and does not reflect the unobserved population, particularly when there are more features than observations (see 4.4.5). Additionally, feature selection may also facilitate a deeper understanding of the scientific question of interest [15]. Feature selection must be performed on the training set to avoid double dipping: the left-out subject features must not be used to determine which features are discriminative. We have followed this approach, as it is clearly described by Ecker et al. [16], and Johnston et al. [17], with two improvements: (i) we used stochastic bootstrapping, running the feature selection heuristic on training set sub-samples and then deciding which features belong to the final “stable” set [18, 19]; (ii) after selecting the the most relevant data, we have employed PCA, obtaining a compact and computationally efficient representation of features in far fewer dimensions [20, 21].

4.4.4 Criteria for the selection of K

We studied how the performance of the classifiers is affected by the selection of the feature set cardinality (K). Figure 4.10 shows that the VBA classifier accuracy is relatively stable across feature set cardinalities. Thus, the optimal K parameter responds to one of two possible research questions: do we want to know the *minimal set* that enables an optimal classification or, do we want to know *all the features* that are relevant to our problem. This topic is treated in depth by Nilson et al. [22], showing that the *ALL-RELEVANT* problem is much harder than the *MINIMAL-OPTIMAL* problem. This dichotomy can also be seen in the context of scientific vs. clinical goals: in machine learning, a minimal feature set that enables optimal classification is ideal; but, when these features describe anatomical landmarks, as in our study, the minimal set may not provide sufficient information of clinical relevance. We proposed a compromise between these two situations by looking at features for a range of different cardinalities. Thus, informative features (Figure 4.11) are defined then as those with high frequency (>80%) among all the “good” reported accuracies (dotted lines in Figure 4.10).

Nonetheless, a systematic exploration of K is not always feasible. L1-penalized logistic regression models can be applied to estimate the *inherent dimensionality* of a binary classification problem. These models set small estimated coefficients to zero accomplishing variable selection [23]. However, the level of regularization determines the sparsity of the solution and the dimensional estimation (number of non-zero coefficients). As an alternative, linear support vector machines can be used to determine K by examining the weights assigned to each dimension by the decision boundary [24, 25]. We adapted the variable selection algorithm described in [18] to estimate an optimal number of features K using a L1-penalized linear SVM. This method adds “dummy” random features to the dataset and then sets an automatic ranking threshold for the “true” features based on the weights assigned to the “dummy” features during training. Methods like this, where features are ranked with an automatic threshold on the ranking criterion are attractive for use in large dimensional spaces [24] and seemed to be adequate for use with our data. This method estimated $K = 264$ for the MSL region, $K = 130$ for the NEO region, and $K = 87$ for the ATL region. Notice that this method falls into the *MINIMAL*

OPTIMAL category.

4.4.5 Limitations

There are three main considerations regarding our results. First, the patient sample is heterogeneous, including lesional (MTS) and non-lesional patients, and due to the reduced sample size we did not perform stratified analyses (controls vs. MTS, controls vs. non-lesional). Nonetheless, as a result of using voxel-based asymmetry, we were able to uniformly analyze patients with left and right seizure onset. Thus, we believe that our approach can be advantageous to other research groups having small patient cohorts.

Second, a reduced sample size is an important consideration in any statistical or automatic classification study. The effect that our sample size has on our MANOVA analysis is appropriately reported in Table 4.3 using partial η^2 and showing a moderate effect. Similarly for support vector machines, a small sample limits the amount of available training data. This scenario is quite common as biological datasets are difficult to acquire, and thus leave-one-out cross-validation is indicated to maximize training and avoid double dipping [2, 15, 17, 21, 26]. It is well known that the prediction error of a classifier can be decomposed in terms of bias (expected error due to model mismatch) and variance (variation due to training sampling and randomization), and that the penalty for using LOOCV is an increased error variance (while error bias is diminished) [27, 28]. For this reason, we advise using linear classifiers under the following circumstances: (i) the sample size is small and (ii) there is a reason to believe that the data can indeed be linearly separable (as per low-dimensional projections, statistical tests, etc). Linear SVMs are advantageous in these cases, since they have a low error variance [29], which can be reduced even further through bagging [30]. In short, the complexity of the model dictates the variability of the decision surface. Therefore, linear classifiers are advisable over non-linear classifiers when training data are limited.

Finally, in many biological data classification scenarios, like ours, it is common to find the so-called *curse of dimensionality*, when there are many more features than observations (subjects). Though it has been shown that support vector machines behave robustly in high

dimensional feature spaces of small datasets compared to other classifiers [31–33], there is a general consensus from previous neuroimaging studies that feature reduction is a fundamental step before applying a predictive model as discussed in the review performed by [15]. Additionally, feature selection is advisable when one of the research objectives is the identification of informative features (as it is our case). To this end, we have implemented the feature selection process as it has been described here, consequently reducing the size of the feature space and mitigating the curse of dimensionality [20].

In summary, a larger patient sample would enable stratified statistical and classification analyses to characterize the asymmetry of individual patient subgroups. Additionally, it would also reduce the amount of cross-validation in the training and feature selection stages.

4.5 Conclusion

This chapter describes a novel approach for the study of neurological conditions characterized by focal asymmetric changes, as it is the case for temporal lobe epilepsy. Our contribution can be summarized by three elements: (i) The proposal of a unified analysis based on the idea of voxel-based asymmetry, independent of the original unit of measurement, therefore leveraging the integration and analysis of multivariate MR data (QR + DTI + GM*); (ii) The identification of individual *asymmetry signatures* that enable both group-level inferences (statistical methods) and individual inferences (individual disease prediction); (iii) The deduction of informative features useful for individual predictions and their verification with respect to previous neuroimaging studies.

Though our sample was relatively small, our results showed a clear distinction between healthy volunteers and patients with TLE in terms of ROI asymmetry in the considered ROIs. Similarly, it was possible to predict the presence of TLE on individual cases with high accuracy (up to 92%). In both the group-level and the individual level analyses, mean diffusivity asymmetry was noteworthy across the temporal lobe, particularly in the hippocampal area. This observation was in agreement with previous DTI studies of TLE showing common patterns of

ipsilateral MD increase in TLE. The relevance of the MD signal was obtained as a result of our integrative approach that allowed the comparison of individual signals in terms of their contribution to the informative feature set. We believe that the multivariate asymmetry analysis has the potential to improve the understanding and diagnosis of temporal lobe epilepsy and similar neurological conditions.

Bibliography

- [1] A. Pepe, I. Dinov, J. Tohka, An automatic framework for quantitative validation of voxel based morphometry measures of anatomical brain asymmetry, *NeuroImage* 100 (C) (2014) 444–459.
- [2] P. Fang, J. An, L.-L. Zeng, H. Shen, F. Chen, W. Wang, S. Qiu, D. Hu, Multivariate pattern analysis reveals anatomical connectivity differences between left and right mesial temporal lobe epilepsy, *NeuroImage: Clinical*.
- [3] J. Zhang, W. Cheng, Z. Wang, Z. Zhang, W. Lu, G. Lu, J. Feng, Pattern classification of large-scale functional brain networks: Identification of informative neuroimaging markers for epilepsy, *PLoS ONE* 7 (5) (2012) e36733.
- [4] S. C. Deoni, High-resolution T1 mapping of the brain at 3T with driven equilibrium single pulse observation of T1 with high-speed incorporation of RF field inhomogeneities (DESPOT1-HIFI), *Journal of Magnetic Resonance Imaging* 26 (4) (2007) 1106–1111.
- [5] M. Jenkinson, S. Smith, A global optimisation method for robust affine registration of brain images, *Medical image analysis* 5 (2) (2001) 143–156.
- [6] H. Huang, C. Ceritoglu, X. Li, A. Qiu, M. I. Miller, P. C. van Zijl, S. Mori, Correction of b0 susceptibility induced distortion in diffusion-weighted images using large-deformation diffeomorphic metric mapping, *Magnetic resonance imaging* 26 (9) (2008) 1294–1302.
- [7] S. C. Deoni, B. K. Rutt, T. M. Peters, Synthetic T1-weighted brain image generation with incorporated coil intensity correction using DESPOT1, *Magnetic resonance imaging* 24 (9) (2006) 1241–1248.
- [8] C. Cortes, V. Vapnik, Support-vector networks, *Mach. Learn.* 20 (3) (1995) 273–297.
- [9] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau,

- M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research* 12 (2011) 2825–2830.
- [10] K. A. Norman, S. M. Polyn, G. J. Detre, J. V. Haxby, Beyond mind-reading: multi-voxel pattern analysis of fMRI data, *Trends in cognitive sciences* 10 (9) (2006) 424–430.
- [11] N. Kriegeskorte, R. Goebel, P. Bandettini, Information-based functional brain mapping, *Proceedings of the National Academy of Sciences of the United States of America* 103 (10) (2006) 3863–3868.
- [12] L. Q. Uddin, V. Menon, C. B. Young, S. Ryali, T. Chen, A. Khouzam, N. J. Minshew, A. Y. Hardan, Multivariate searchlight classification of structural magnetic resonance imaging in children and adolescents with autism, *Biological psychiatry* 70 (9) (2011) 833–841.
- [13] Z. Yang, F. Fang, X. Weng, Recent developments in multivariate pattern analysis for functional MRI, *Neuroscience bulletin* 28 (4) (2012) 399–408.
- [14] N. K. Focke, M. Yogarajah, M. R. Symms, O. Gruber, W. Paulus, J. S. Duncan, Automated MR image classification in temporal lobe epilepsy, *NeuroImage* 59 (1) (2012) 356–362.
- [15] B. Mwangi, T. Tian, J. Soares, A review of feature reduction techniques in neuroimaging, *Neuroinformatics* (2013) 1–16.
- [16] C. Ecker, V. Rocha-Rego, P. Johnston, J. Mourao-Miranda, A. Marquand, E. M. Daly, M. J. Brammer, C. Murphy, D. G. Murphy, M. A. Consortium, et al., Investigating the predictive value of whole-brain structural mr scans in autism: a pattern classification approach, *Neuroimage* 49 (1) (2010) 44–56.
- [17] B. Johnston, B. Mwangi, K. Matthews, D. Coghill, J. Steele, Predictive classification of individual magnetic resonance imaging scans from children and adolescents, *European child & adolescent psychiatry* 22 (12) (2013) 733–744.

- [18] J. Bi, K. Bennett, M. Embrechts, C. Breneman, M. Song, Dimensionality reduction via sparse support vector machines, *The Journal of Machine Learning Research* 3 (2003) 1229–1243.
- [19] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, *J. Mach. Learn. Res.* 3 (2003) 1157–1182.
- [20] A. Golugula, G. Lee, A. Madabhushi, Evaluating feature selection strategies for high dimensional, small sample size datasets, in: *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE, IEEE, 2011*, pp. 949–952.
- [21] D. Cantor-Rivera, A. R. Khan, M. Goubran, S. M. Mirsattari, T. M. Peters, Detection of temporal lobe epilepsy using support vector machines in multi-parametric quantitative mr imaging, *Computerized Medical Imaging and Graphics*.
- [22] R. Nilsson, J. M. Peña, J. Björkegren, J. Tegnér, Consistent feature selection for pattern recognition in polynomial time, *The Journal of Machine Learning Research* 8 (2007) 589–612.
- [23] J. Fan, R. Li, Statistical challenges with high dimensionality: Feature selection in knowledge discovery, *arXiv preprint math/0602133*.
- [24] I. Guyon, J. Weston, S. Barnhill, V. Vapnik, Gene selection for cancer classification using support vector machines, *Machine learning* 46 (1-3) (2002) 389–422.
- [25] Y.-W. Chang, C.-J. Lin, Feature ranking using linear svm, *Causation and Prediction Challenge Challenges in Machine Learning* 2 (2008) 47.
- [26] D. H. Foley, Considerations of sample and feature size, *Information Theory, IEEE Transactions on* 18 (5) (1972) 618–626.
- [27] B. Clarke, E. Fokoue, H. H. Zhang, *Principles and theory for data mining and machine learning*, Springer Science & Business Media, 2009.

- [28] T. Hastie, R. Tibshirani, J. Friedman, T. Hastie, J. Friedman, R. Tibshirani, The elements of statistical learning, Vol. 2, Springer, 2009.
- [29] C. D. Manning, P. Raghavan, H. Schütze, Introduction to information retrieval, Vol. 1, Cambridge university press Cambridge, 2008.
- [30] L. Breiman, Bias, variance, and arcing classifiers, Statistics.
URL <http://oz.berkeley.edu/~breiman/arcall96.pdf>
- [31] J. Hua, Z. Xiong, J. Lowey, E. Suh, E. R. Dougherty, Optimal number of features as a function of sample size for various classification rules, *Bioinformatics* 21 (8) (2005) 1509–1515.
- [32] Y. Guo, A. Graber, R. N. McBurney, R. Balasubramanian, Sample size and statistical power considerations in high-dimensionality data settings: a comparative study of classification algorithms, *BMC bioinformatics* 11 (1) (2010) 447.
- [33] M. C. Alonso, J. A. Malpica, A. M. de Agirre, Consequences of the hughes phenomenon on some classification techniques, in: ASPRS 2011 Annual Conference, Milwaukee, Wisconsin May, 2011, pp. 1–5.

Chapter 5

Conclusions

5.1 Contributions

This thesis provided a conceptual framework for integrating quantitative relaxometry and DTI into TLE diagnostic application. In Chapter 2, feature maps were obtained for every subject showing distinct abnormality patterns for MTS+ and MTS- patients (Sections 2.3.1 and 2.3.2); a *laterality score* evaluated on these maps obtained a good correspondence between the predicted side of seizure onset and the surgery performed (ground truth). Similarly, Chapters 3 and 4 demonstrated how Q-MRI and DTI can be combined to predict a TLE diagnosis, using ROI- and voxel-level features respectively. Support vector machines showed optimal diagnostic accuracy accounting for our small and heterogeneous patient group.

Consequently, the analysis of informative features presented herein provided clues for structural abnormalities in TLE patient group as well as in the examined patient subgroups. By virtue of the multi-parametric evaluation, it was possible to determine the relevance of mean diffusivity relative to other signals. Hence, future studies could focus on this type of image. Similarly, feature analysis revealed that most of the abnormalities in the MTS- group are located in the temporal cortex and are significant compared to controls (Figure 2.7). This is an interesting finding since non-MTS patients have a worse surgical outcome than MTS patients (50% vs 80%) [1]. Hence, a more detailed examination of cortical features in this group has

the potential of identify lesions, improving the patient's odds of becoming seizure-free after surgery.

Another contribution that this work makes is the examination of machine learning classification models of TLE. This analysis demonstrated that the image type is more significant than the type of SVM, with MD images being the most relevant in classification using most models. Additionally, we showed that linear SVMs are sufficient to identify TLE patients and should be preferred over non-linear models that require more extensive parameter optimization, which is limiting given a small sample. Also, patients with right TLE were more difficult to detect, and in these cases the multi-image classification proved to be optimal.

This work demonstrated the fitness of using a two-sample Kolmogorov-Smirnov test to assess asymmetry in regions of interest and in voxel-level features. This test is non-parametric and thus befits left-right comparisons of regions that depart from normality due to the presence of pathologies. Additionally, the KS test standardizes left-right differences to the [0,1] range and this allows the comparison of asymmetry across different imaging techniques such as those in the scope of this thesis. The progression from ROI to voxel-based analyses demonstrated that classification is optimal even when restricted to the temporal lobe. Moreover, interesting asymmetries in the hippocampus, amygdala, parahippocampal gyrus (entorhinal and perirhinal cortices), and distinctive white matter were identified without previous segmentation of these structures, with no involvement of the medial and superior temporal cortices (Figure 4.11).

In practice, adding Q-MRI and DTI to clinical protocols for the evaluation of TLE may be warranted when current methods fail to identify epileptogenic foci. In these scenarios, the methods described in this thesis have the potential to provide valuable diagnostic information prior or as an alternative to intracranial EEG. A first and simple approach for integrating this thesis in clinical workflow could be the evaluation of MD images with an ANOVA-based linear-SVM model, and the evaluation of IFMs in the images that are currently used for diagnosis.

5.2 Limitations

The main constraint during the development of the experiments described in this thesis was the small patient group. This limitation is well known in neuroscience and neuroimaging studies (Section 1.4.1). In this respect, support vector machines are a relatively robust technique in comparison to other classifiers like random forests, linear discriminant analysis and neural networks [2, 3]. However, how can we determine if there are enough data for obtaining a reliable boundary? A possible approach to solve this question was discussed in Section 3.3.6 in which the SVM reliability index (SRI) measures the variability of the decision boundary when data points are removed. Notice that this idea is linked to cross-validation, where data points are also removed before estimating the decision boundary. Hence, SRI could serve as an upper boundary for determining the optimal partition sizes in for leave-n-out cross-validation.

Another limitation is the effect that cross-validation has in the identification of informative features. Feature selection must be stratified, evaluating features on every possible training set created with cross-validation. Then heuristics like those presented in Sections 3.2.8 and 4.2.5 are needed to assess valuable features. The problem with this approach is that new parameters are being added to the problem: how do we decide what a good frequency threshold is so we can call a feature *informative*? Different thresholds will vary the number of features identified and in this thesis these thresholds have been arbitrarily set to be in the 80% (Chapter 3) and 100% (Chapter 4).

Similarly, feature selection methods add an extra cardinality parameter, K , and though it is interesting to test the robustness of SVMs by restricting the number of features, a complete evaluation of all possible K is not always feasible or practical (Section 4.4.4). As discussed in this thesis, L1-penalized regression models can be used to estimate K but, these models depend on the level of regularization chosen. Though this regularization can be estimated by cross-validation, in my view, this is not an optimal approach. Hence, the exploration of *embedded* methods for feature selection is an interesting venue for future research. As a matter of fact, a solution to the cardinality dilemma can be obtained by applying “*Dimensionality reduction via*

sparse support vector machines” as described by Bi et al. [4] to estimate an optimal number of features K using a L1-penalized linear SVM (Section 4.4.4).

5.3 Future Research

A possible future area of research is the comparison of IFM results with inter-ictal EEG studies. Region correspondences could benefit EEG-based diagnosis and also provide a heuristic for better electrode positioning. Correspondingly, the voxel-based asymmetry maps obtained in Chapter 4 could accomplish a similar purpose with respect to intracranial EEG due to their higher spatial resolution (compared to the ROI approach). Additionally, white matter tractography could serve as *a priori* knowledge source for linking abnormal ROIs identified with IFM to EEG studies. However, as shown in Figure 1.14 and discussed in previous works (VBR vs. ROI [5]; TBSS vs. VBM [6]; and TBSS vs. VBS [7]) multiple techniques may have a complementary character rather than an additive effect. Thus, established techniques must not be taken as a gold standard, since each technique has an inherent bias, sensitivity and specificity.

Another venue for research is to focus on detection and lateralization of non-MTS patients since this group has a higher morbidity (including patients with FCD). To this end, our results indicate that the exploration of temporal cortical abnormalities may be useful. This topic has previously explored through cortical thickness estimation in MRI [8], and stereology of surgical specimens [9].

In regards to SVMs in a clinical perspective, larger patient samples would allow to investigate the correlation between subject-to-boundary signed distances and clinical variables as shown by Ecker et al.(2010) [10] for Autism (Figure 5.1). This was evaluated in our study but no statistical significance was obtained. From a scientific perspective, it would be interesting to explore *deep learning* algorithms for obtaining high-level representations of features. These methods are unsupervised, can boost SVM-based classification, and have gathered interest largely due to recent developments [11, 12].

Finally, though a significant pattern of increasing asymmetry was observed among controls,

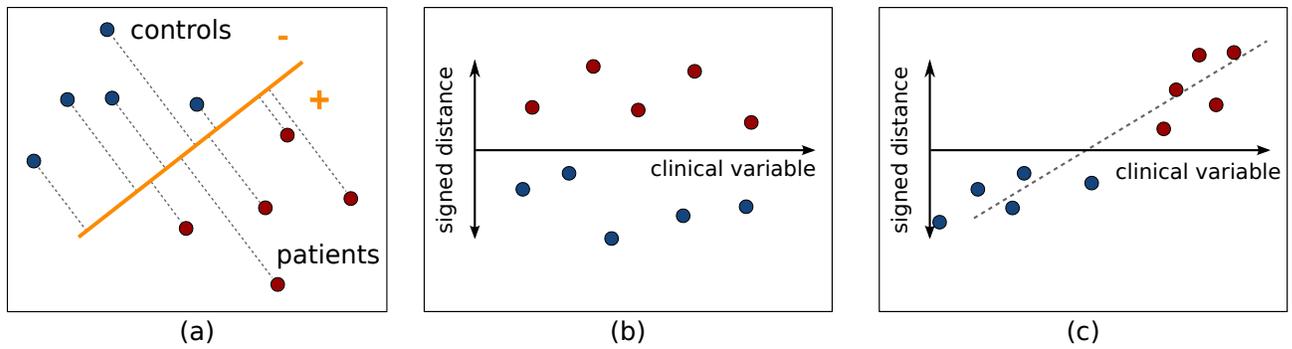


Figure 5.1: *Subject-to-boundary correlation*. The idea is to correlate the signed distance from each subject to the decision boundary shown in (a) with clinical variables. In other words, can we associate distance to the severity of the clinical condition being examined? (b) shows a case where there is no correlation; and (c) a case where a correlation has been found.

MTS- and MTS+ groups (Figure 2.8a), a stratified analysis of the asymmetry was not possible using voxel-based asymmetry (e.g., left-non-MTS vs. right-MTS). Figure 5.2 shows some preliminary results. A larger patient cohort would warrant analysis at this level.

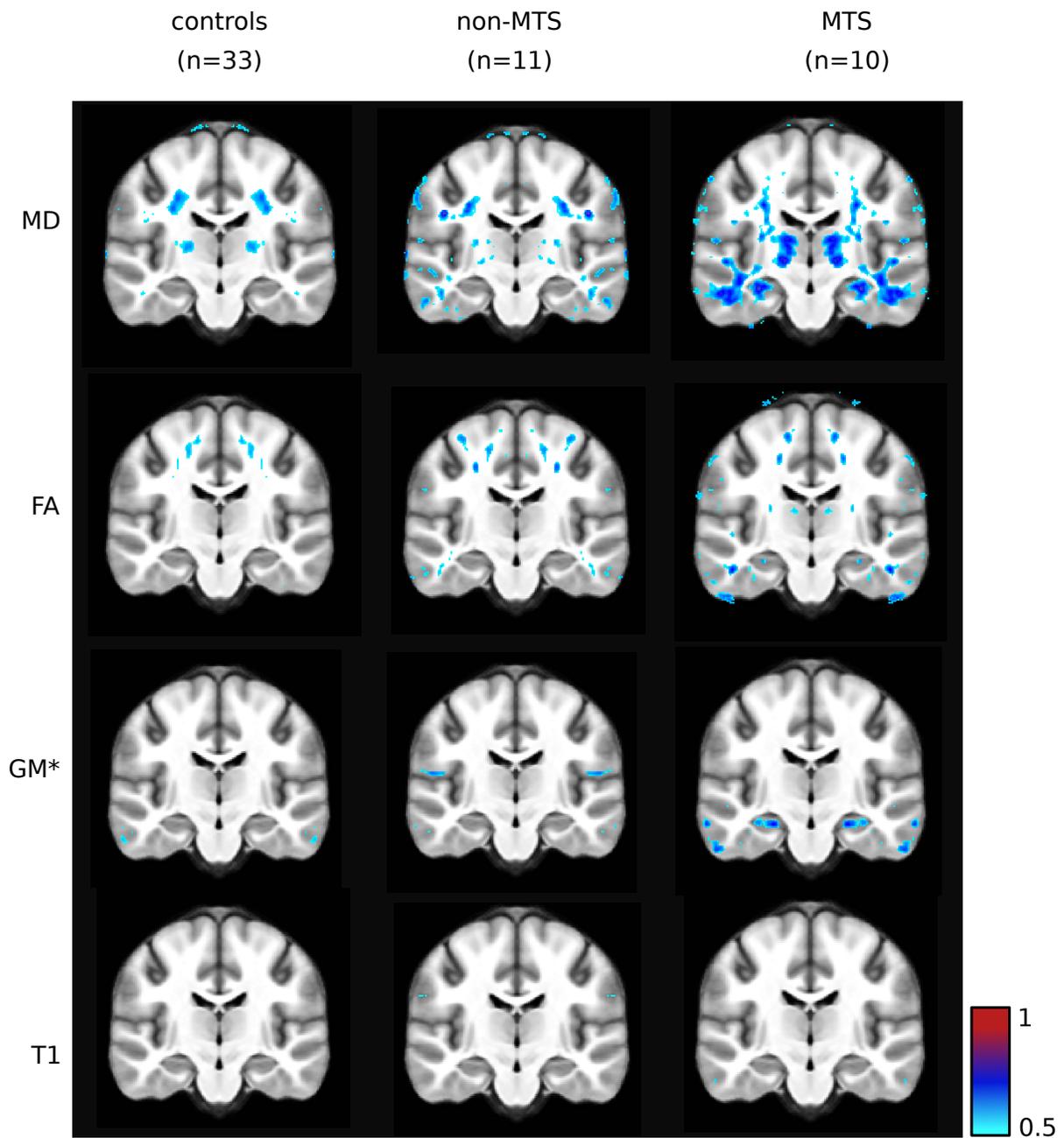


Figure 5.2: **voxel-based asymmetry patterns.** Each row shows an average asymmetry map. The asymmetry threshold is set at 50%

Bibliography

- [1] V. Gupta, R. Bronen A., Epilepsy, in: S. W. Atlas (Ed.), *Magnetic resonance imaging of the brain and spine*, Vol. 1, Lippincott Williams & Wilkins, 2009, pp. 307–342.
- [2] J. Hua, Z. Xiong, J. Lowey, E. Suh, E. R. Dougherty, Optimal number of features as a function of sample size for various classification rules, *Bioinformatics* 21 (8) (2005) 1509–1515.
- [3] Y. Guo, A. Graber, R. N. McBurney, R. Balasubramanian, Sample size and statistical power considerations in high-dimensionality data settings: a comparative study of classification algorithms, *BMC bioinformatics* 11 (1) (2010) 447.
- [4] J. Bi, K. Bennett, M. Embrechts, C. Breneman, M. Song, Dimensionality reduction via sparse support vector machines, *The Journal of Machine Learning Research* 3 (2003) 1229–1243.
- [5] G. S. Pell, R. S. Briellmann, H. Pardoe, D. F. Abbott, G. D. Jackson, Composite voxel-based analysis of volume and t2 relaxometry in temporal lobe epilepsy, *Neuroimage* 39 (3) (2008) 1151–1161.
- [6] C. Scanlon, S. G. Mueller, I. Cheong, M. Hartig, M. W. Weiner, K. D. Laxer, Grey and white matter abnormalities in temporal lobe epilepsy with and without mesial temporal sclerosis, *Journal of Neurology* 260 (9) (2013) 2320–2329.
- [7] N. K. Focke, M. R. Symms, J. L. Burdett, J. S. Duncan, Voxel-based analysis of whole brain flair at 3t detects focal cortical dysplasia, *Epilepsia* 49 (5) (2008) 786–793.
- [8] C. R. McDonald, D. J. Hagler, M. E. Ahmadi, E. Tecoma, V. Iragui, L. Gharapetian, A. M. Dale, E. Halgren, Regional neocortical thinning in mesial temporal lobe epilepsy, *Epilepsia* 49 (5) (2008) 794–803.
- [9] S. Bothwell, G. E. Meredith, J. Phillips, H. Staunton, C. Doherty, E. Grigorenko, S. Glazier, S. A. Deadwyler, C. A. O’Donovan, M. Farrell, Neuronal hypertrophy in the

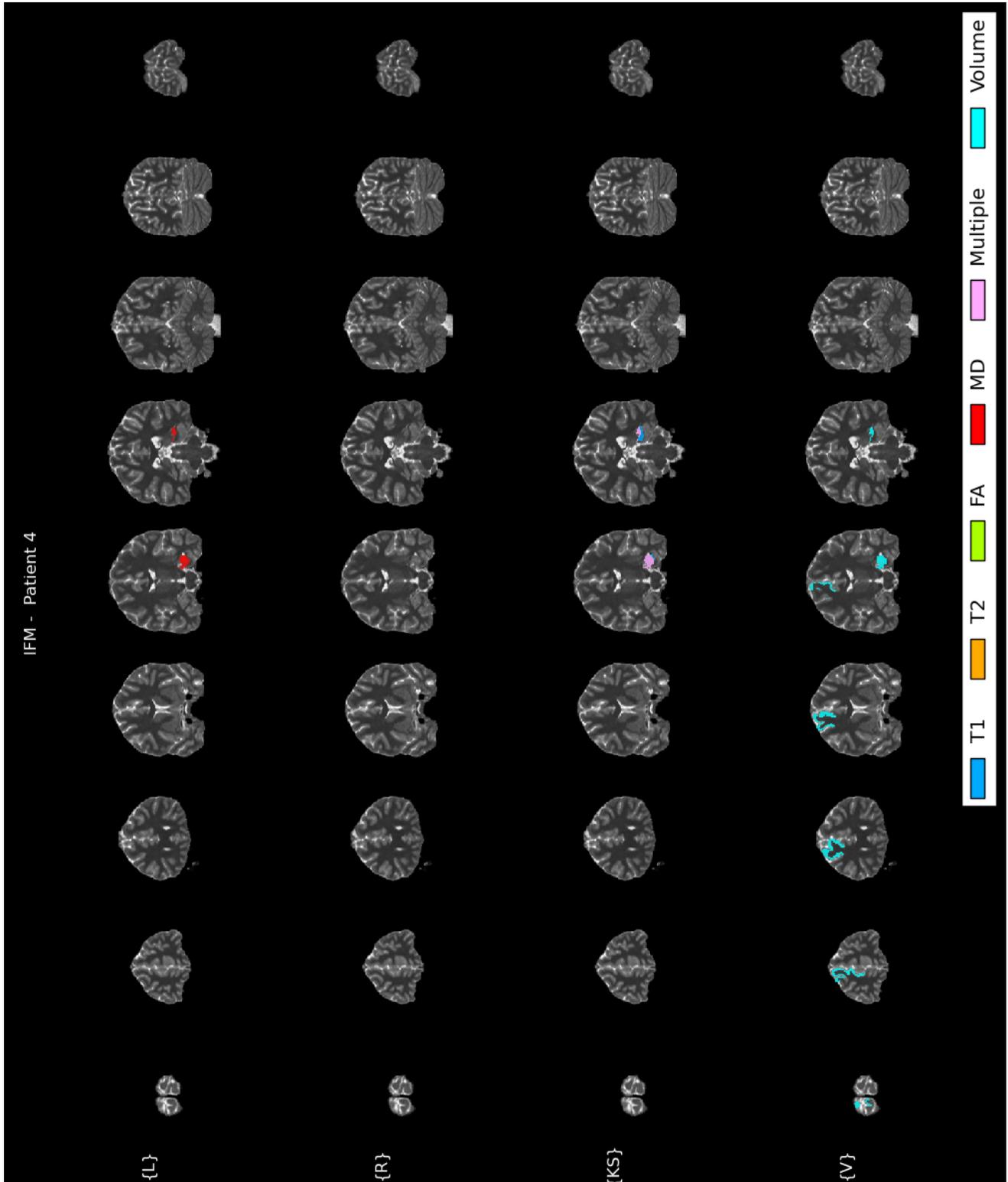
neocortex of patients with temporal lobe epilepsy, *The Journal of Neuroscience* 21 (13) (2001) 4789–4800.

- [10] C. Ecker, V. Rocha-Rego, P. Johnston, J. Mourao-Miranda, A. Marquand, E. M. Daly, M. J. Brammer, C. Murphy, D. G. Murphy, M. A. Consortium, et al., Investigating the predictive value of whole-brain structural mr scans in autism: a pattern classification approach, *Neuroimage* 49 (1) (2010) 44–56.
- [11] G. Hinton, S. Osindero, Y.-W. Teh, A fast learning algorithm for deep belief nets, *Neural computation* 18 (7) (2006) 1527–1554.
- [12] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P.-A. Manzagol, Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion, *The Journal of Machine Learning Research* 11 (2010) 3371–3408.

Appendix A

Individual Feature Maps

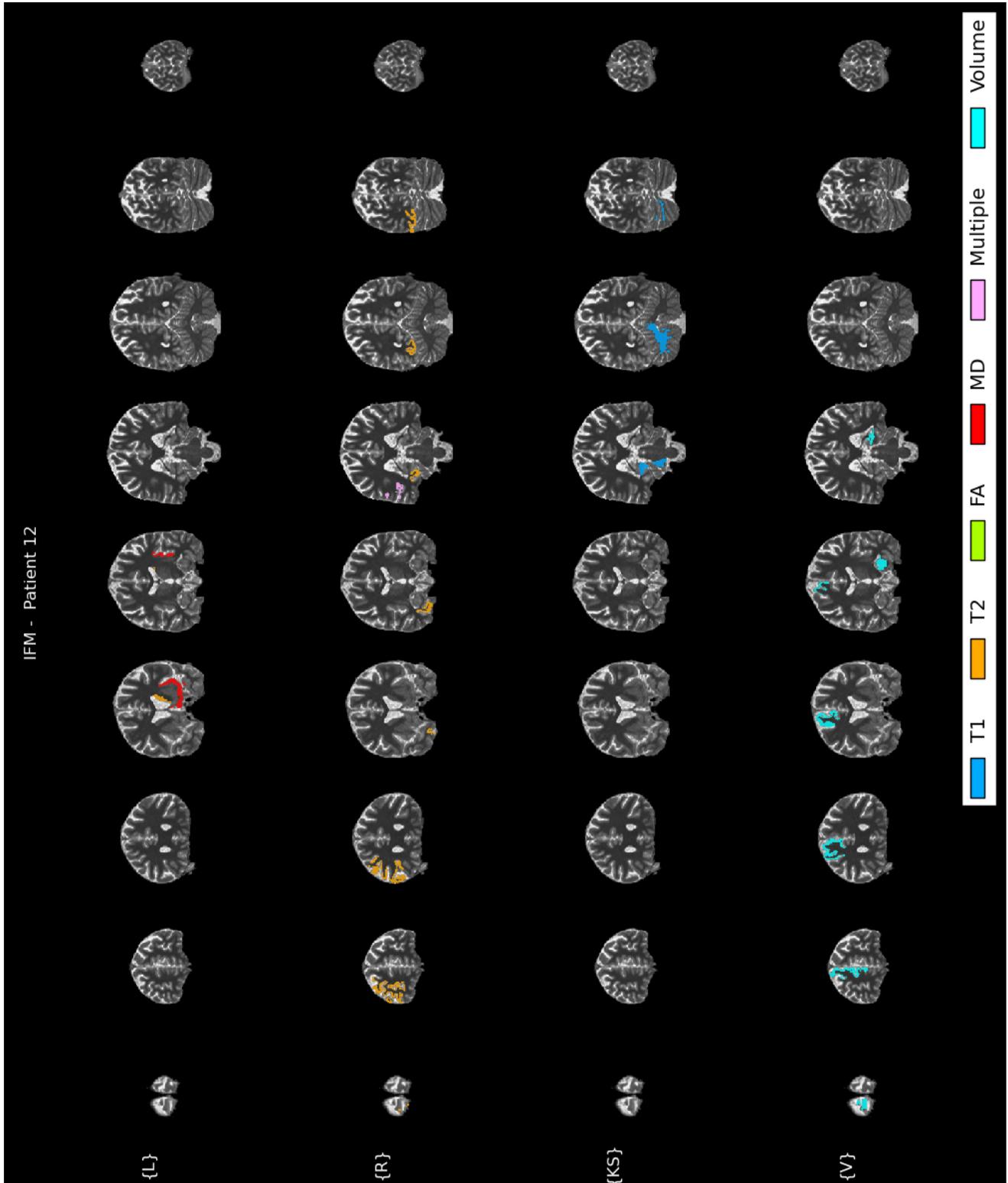
A.1 IFM for a MTS patient



```
1 Individual Feature Maps (IFM) - Report
2 Patient: 4
3 =====
4
5 LATERALITY SCORE: -1.00 [-1: left, +1: right]
6
7 Left Mean Intensity Map {L}
8 --
9 T1
10 --
11 * NO FINDINGS
12 --
13 T2
14 --
15 * NO FINDINGS
16 --
17 FA
18 --
19 * NO FINDINGS
20 --
21 MD
22 --
23     Left-Hippocampus (temporal)
24         Higher 0.00113 [0.00096]
25
26 Right Mean Intensity Map {R}
27 --
28 T1
29 --
30 * NO FINDINGS
31 --
32 T2
33 --
34 * NO FINDINGS
35 --
36 FA
37 --
38 * NO FINDINGS
39 --
40 MD
41 --
42 * NO FINDINGS
43 Signal Asymmetry {KS}
44 --
45 T1
46 --
47     wm-lh-parahippocampal vs wm-rh-parahippocampal (temporal)
48         Abnormal 0.20 [0.06]
49         Left Higher 1400.64(L) > 1313.74(R)
50 --
51 T2
52 --
53     Left-Hippocampus vs Right-Hippocampus (temporal)
54         Abnormal 0.35 [0.09]
55         Left Higher 72.02(L) > 55.23(R)
56 --
```

57 FA
58 --
59 * NO FINDINGS
60 --
61 MD
62 --
63 Left-Hippocampus vs Right-Hippocampus (temporal)
Abnormal 0.53136 [0.09038]
64 Left Higher 0.00113(L) > 0.00093(R)
65
66
67 Volume Asymmetry {V}
68
69 ctx-lh-superiorfrontal vs ctx-rh-superiorfrontal (frontal)
Abnormal 0.003663 [0.000807]
70 Right Smaller 0.026560(L) > 0.022897(R)
71
72 Left-Hippocampus vs Right-Hippocampus (temporal)
Abnormal 0.001014 [0.000154]
73 Left Smaller 0.003475(L) < 0.004490(R)

A.2 IFM for a non-MTS patient



```

1 Individual Feature Maps (IFM) - Report
2 Patient: 12
3 =====
4
5 LATERALITY SCORE: 1.00 [-1: left, +1: right]
6
7
8 Left Mean Intensity Map {L}
9 --
10 T1
11 --
12 * NO FINDINGS
13 --
14 T2
15 --
16     Left-Caudate (other)
17         Higher 83.25 [44.04]
18 --
19 FA
20 --
21 * NO FINDINGS
22 --
23 MD
24 --
25     ctx-lh-insula (other)
26         Higher 0.00108 [0.00091]
27 --
28 Right Mean Intensity Map {R}
29 --
30 T1
31 --
32 * NO FINDINGS
33 --
34 T2
35 --
36     ctx-rh-parstriangularis (frontal)
37         Higher 86.90 [57.90]
38     ctx-rh-bankssts (temporal)
39         Higher 65.09 [46.05]
40     ctx-rh-rostralmiddlefrontal (frontal)
41         Higher 85.14 [57.50]
42 --
43 FA
44 --
45 * NO FINDINGS
46 --
47 MD
48 --
49     ctx-rh-bankssts (temporal)
50         Higher 0.00100 [0.00084]
51     ctx-rh-fusiform (temporal)
52         Higher 0.00091 [0.00082]
53 --
54 Signal Asymmetry {KS}
55 --
56 T1
57 --
58     ctx-lh-parahippocampal vs ctx-rh-parahippocampal (temporal)
59         Abnormal 0.16 [0.05]

```

```
52      Right Higher 1965.75(L) < 2053.67(R)
53
54      wm-lh-parahippocampal vs wm-rh-parahippocampal (temporal)
          Abnormal 0.18 [0.06]
55      Right Higher 1334.48(L) < 1403.61(R)
56
57      Left-Cerebellum-White-Matter vs Right-Cerebellum-W (other)
          Abnormal 0.19 [0.05]
58      Right Higher 1213.76(L) < 1255.19(R)
59
60  --
61  T2
62  --
63  * NO FINDINGS
64  --
65  FA
66  --
67  * NO FINDINGS
68  --
69  MD
70  --
71  * NO FINDINGS
72
73  Volume Asymmetry {V}
74
75      ctx-lh-superiorfrontal vs ctx-rh-superiorfrontal (frontal)
          Abnormal 0.003634 [0.000807]
76      Right Smaller 0.023142(L) > 0.019508(R)
77
78      Left-Hippocampus vs Right-Hippocampus (temporal)
          Abnormal 0.000788 [0.000154]
79      Left Smaller 0.003307(L) < 0.004095(R)
```

Appendix B

Research Ethics Board Approval



Office of Research Ethics

The University of Western Ontario

Website: www.uwo.ca/research/ethics

Use of Human Subjects - Ethics Approval Notice

Principal Investigator: Dr. T.M. Peters

Review Number: 16189

Review Level: Full Board

Review Date: May 19, 2009

Protocol Title: Structural and Functional MR imaging in Frontal and Temporal Lobe Epilepsy at 1.5T, 3T, and 7T

Department and Institution: Imaging, Robarts Research Institute

Sponsor: CIHR-CANADIAN INSTITUTE OF HEALTH RESEARCH

Ethics Approval Date: October 7, 2009

Expiry Date: July 31, 2015

Documents Reviewed and Approved: UWO Protocol, Letter of information & consent form dated Aug. 31/09 & Advertisement dated Aug. 31/09

Documents Received for Information:

This is to notify you that The University of Western Ontario Research Ethics Board for Health Sciences Research Involving Human Subjects (HSREB) which is organized and operates according to the Tri-Council Policy Statement: Ethical Conduct of Research Involving Humans and the Health Canada/ICH Good Clinical Practice Practices: Consolidated Guidelines; and the applicable laws and regulations of Ontario has reviewed and granted approval to the above referenced study on the approval date noted above. The membership of this REB also complies with the membership requirements for REB's as defined in Division 5 of the Food and Drug Regulations.

The ethics approval for this study shall remain valid until the expiry date noted above assuming timely and acceptable responses to the HSREB's periodic requests for surveillance and monitoring information. If you require an updated approval notice prior to that time you must request it using the UWO Updated Approval Request Form.

During the course of the research, no deviations from, or changes to, the protocol or consent form may be initiated without prior written approval from the HSREB except when necessary to eliminate immediate hazards to the subject or when the change(s) involve only logistical or administrative aspects of the study (e.g. change of monitor, telephone number). Expedited review of minor change(s) in ongoing studies will be considered. Subjects must receive a copy of the signed information/consent documentation.

Investigators must promptly also report to the HSREB:

- a) changes increasing the risk to the participant(s) and/or affecting significantly the conduct of the study;
- b) all adverse and unexpected experiences or events that are both serious and unexpected;
- c) new information that may adversely affect the safety of the subjects or the conduct of the study.

If these changes/adverse events require a change to the information/consent documentation, and/or recruitment advertisement, the newly revised information/consent documentation, and/or advertisement, must be submitted to this office for approval.

Members of the HSREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the HSREB.

Original signed by Janice Sutherland

Chair of HSREB: Dr. Joseph Gilbert

Ethics Officer to Contact for Further Information			
<input checked="" type="checkbox"/> Janice Sutherland	<input type="checkbox"/> Elizabeth Wambolt	<input type="checkbox"/> Grace Kelly	<input type="checkbox"/> Denise Grafton

This is an official document. Please retain the original in your files.

cc: ORE File
LHRI

CURRICULUM VITAE

Name Diego Hernando Cantor Rivera

Education

2015 Ph.D. Candidate in Biomedical Engineering
Robarts Research Institute, Western University, London, ON

2008 M.Eng. in Computer and Systems Engineering
Universidad de Los Andes, Bogotá, Colombia

2001 B.Eng. in Design and Automation Engineering
Universidad de La Salle, Bogotá, Colombia

Honours and Awards:

2009-2010 Computer Assisted Medical Intervention (CAMI)
Training Program Scholarship (CAD \$10000/yr)

2008-2013 Western Graduate Research Scholarship (WGRS) - (CAD \$17500/yr)

2007 “Gonzalo Esguerra” Award to Academic Excellence
given by the Colombian Society of Radiology

Work Experience

2008 Industrial Trainee
CSIRO Australian e-Health Research Centre
Brisbane, Australia

2007 Stagiare
CREATIS-LRMN
INSA, Lyon, France

2005 Information Associate
EDS Electronic Data Systems (acquired by Hewlett-Packard)
Bogotá, Colombia

2003 Project Leader
Informática Siglo XXI
Bogotá, Colombia

Teaching

2009-2010 Teaching Assistant, Western University, London, ON
Faculty of Engineering, Biomedical Engineering Graduate Program
Introduction to Medical Imaging

2009 Teaching Assistant, Western University, London, ON
Faculty of Engineering, Dept. of Electrical and Computer Engineering
Programming Fundamentals for Engineers

2006-2007 Research Assistant, Universidad de Los Andes, Bogotá, Colombia
Faculty of Engineering, Dept. of Computer and Systems Engineering
Digital Image Processing

2005 Instructor, Universidad de Los Andes, Bogotá, Colombia
Faculty of Engineering, Dept. of Computer and Systems Engineering
Object Oriented Programming Fundamentals (APO1)

Invited Talks

2015 *Machine Learning and MRI imaging for Temporal Lobe Epilepsy*
In: epilepsy rounds seminar
Dr. Jorge Burneo, London Health Sciences Centre, London, ON

2014 *Introduction to WebGL programming*
In: Software Engineering for Human-Computer Interface Design course
Prof. Roy Eagleson, Western University, London, ON

2013 *Rendering virtual medical models on the web in real-time*
In: IMAGINE research group seminar on Biomedical Imaging
Prof. Marcela Hernandez-Hoyos, Universidad de Los Andes, Bogotá,
Colombia

2007 *Patient-specific MRI-based vascular model generation using fast marching*
In: II Seminar on Mathematical Modeling of Medical Images, Bogotá,
Colombia
In: IV Seminar of Biomedical Engineering, Bogotá, Colombia

Books and chapters

1. **D. Cantor**, B. Jones, “*WebGL Beginner’s Guide*”, Packt Publishing (2012).
2. **D. Cantor**, K. Abhari “Creating cameras in WebGL: current challenges and possibilities”, In: P. Cozzi editor. *WebGL Insights. SIGGRAPH2015*, CRC Press (2015).

Journal Publications

1. **D. Cantor-Rivera**, A.R. Khan, M. Goubran, S.M. Mirsattari, T.M. Peters, “*Detection of temporal lobe epilepsy using support vector machines in multi-parametric quantitative MR imaging*”, *Computerized Medical Imaging and Graphics* 41 (2015) 14-28.
2. **D. Cantor-Rivera**, S. de Ribaupierre, J.C. Lau, J.S.H. Baxter, S.M. Mirsattari, J.G. Burneo, D.A. Steven, T.M. Peters, A.R. Khan, “*Automated classification of temporal lobe epilepsy patients using voxel-based asymmetry (VBA)*”, *NeuroImage: Clinical* (submitted).
3. **D. Cantor-Rivera**, S. de Ribaupierre, J.C. Lau, S. Mirsattari, M. Goubran, J.G. Burneo, D.A. Steven, T.M. Peters, A.R. Khan, “*Individual Feature Maps: a patient-specific analysis tool with applications in temporal lobe epilepsy*”, *International Journal of Computer Assisted Radiology and Surgery* (submitted).
4. **D. Cantor**, M. Orkisz, J. Arias, L.F. Uriza, M. Hernández Hoyos, “*Modelamiento en 3D de patologías vasculares a partir de imágenes de resonancia magnética sustrada*” [3D Modeling of vascular pathologies from contrast enhanced magnetic resonance images (MRI)], *Rev. Colomb. Radiol* 18(4):2225-2232 (2007).

Conference Publications

1. **D. Cantor-Rivera**, R. Bartha, T. Peters, *Efficient 3D rendering for Web-based Medical Imaging Software: a proof of concept*, Proc. SPIE 7964, 79643A (2011); doi:10.1117/12.878814
2. **D. Cantor-Rivera**, T. Peters, *Pervasive Medical Imaging Applications: current challenges and possible alternatives*, mHealth Workshop - IEEE eHealthComm2010 (2010), Lyon, France.
3. **D. Cantor-Rivera**, M. Goubran, A. Kraguljac, R. Bartha, T. Peters, *Detection of small human cerebral cortical lesions with MRI under different levels of Gaussian smoothing: applications in epilepsy*, Proc. SPIE 7623, 76233S (2010); doi:10.1117/12.845058

4. **D. Cantor-Rivera** , M. Hernandez-Hoyos, *Generacin de modelos arteriales tridimensionales a partir de imgenes de resonancia magntica (IRM)* [Tridimensional artery model generation based on magnetic resonance imaging MRI], Conferencia Latinoamericana de Informtica (CLEI2007) (2007), Costa Rica.
5. S. Rojas, **D. Cantor** , L.D. Rubiano, P. Melo, C. Patino, *Using Human Gaze as Input Data for Computer Programs*, International Conference on Computer Science, Software Engineering, Information Technology, e-Business, and Applications (CSITeA '02) (2002), Brazil.
6. L.D. Rubiano, **D. Cantor**, S. Rojas, ”*Eye Tracking and Gaze Estimation using JMF and IMAQ Vision*”, Visualization, Imaging, and Image Processing (VIIP 2002) (2002), Spain.

Poster Presentations

- | | |
|------|--|
| 2015 | International Society for Magnetic Resonance in Medicine (ISMRM), Toronto, ON |
| 2014 | London Imaging Day, London, ON |
| 2014 | Southern Ontario Neuroscience Asociation (SONA), London, ON |
| 2013 | Graphics, Animation and New Media (GRAND), Toronto, ON |
| 2013 | International Society for Magnetic Resonance in Medicine (ISMRM), Salt Lake City, UT |
| 2011 | Medical Image Computing and Computer Assisted Intervention (MICCAI), Toronto, ON |
| 2011 | SPIE Medical Imaging conference, Orlando, FL |
| 2010 | International Society for Magnetic Resonance in Medicine (ISMRM), Stockholm, Sweden |
| 2010 | Imaging Network Ontario (ImNO), Toronto, ON |
| 2009 | London Imaging Day, London, ON |
| 2007 | Medical Imaging Research Day, Bogotá, Colombia |