



HAL
open science

Classification and statistical modeling based on clinical and conventional and advanced Magnetic Resonance Imaging data

Ceren Tozlu

► **To cite this version:**

Ceren Tozlu. Classification and statistical modeling based on clinical and conventional and advanced Magnetic Resonance Imaging data. Bioinformatics [q-bio.QM]. Université de Lyon, 2018. English. NNT: 2018LYSE1043 . tel-01881783

HAL Id: tel-01881783

<https://theses.hal.science/tel-01881783>

Submitted on 26 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N° d'ordre NNT : xxx

THÈSE DE DOCTORAT DE L'UNIVERSITÉ DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

École Doctorale: Ecosystèmes Evolution Modélisation Microbiologie (E2M2)
Laboratoire de Biométrie et Biologie Évolutive UMR CNRS 5558

Spécialité de doctorat : Biostatistiques, Santé Publique

Soutenue publiquement le 14/03/2018, par :
Ceren Tozlu

**Classification et modélisation statistique intégrant des
données cliniques et d'imagerie par résonance magnétique
conventionnelle et avancée**

Devant le jury composé de :

Bekiroglu Nural , PU, <i>Marmara University</i>	Rapporteuse
Felblinger Jacques , PU-PH, <i>Université de Lorraine</i>	Rapporteur
Hommel Marc , PU-PH, <i>Université Grenoble Alpes</i>	Rapporteur
Durand-Dubief Françoise , PH, <i>Hospices Civils de Lyon</i>	Examinatrice
Ouchchane Lemlih , MCU-PH, <i>Université Clermont-Ferrand 1</i>	Examinateur
Maucort-Boulch Delphine , PU-PH, <i>Université de Lyon</i>	Directrice de thèse
Sappey-Marinier Dominique , MCU-PH, <i>Université de Lyon</i>	Co-directeur de thèse

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université

Président du Conseil Académique

Vice-président du Conseil d'Administration

Vice-président du Conseil Formation et Vie Universitaire

Vice-président de la Commission Recherche

Directrice Générale des Services

M. le Professeur Frédéric FLEURY

M. le Professeur Hamda BEN HADID

M. le Professeur Didier REVEL

M. le Professeur Philippe CHEVALIER

M. Fabrice VALLÉE

Mme Dominique MARCHAND

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard

Faculté de Médecine et de Maïeutique Lyon Sud – Charles
Mérieux

Faculté d'Odontologie

Institut des Sciences Pharmaceutiques et Biologiques

Institut des Sciences et Techniques de la Réadaptation

Département de formation et Centre de Recherche en Biologie
Humaine

Directeur : M. le Professeur G.RODE

Directeur : Mme la Professeure C. BURILLON

Directeur : M. le Professeur D. BOURGEOIS

Directeur : Mme la Professeure C. VINCIGUERRA

Directeur : M. X. PERROT

Directeur : Mme la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies

Département Biologie

Département Chimie Biochimie

Département GEP

Département Informatique

Département Mathématiques

Département Mécanique

Département Physique

UFR Sciences et Techniques des Activités Physiques et Sportives

Observatoire des Sciences de l'Univers de Lyon

Polytech Lyon

Ecole Supérieure de Chimie Physique Electronique

Institut Universitaire de Technologie de Lyon 1

Ecole Supérieure du Professorat et de l'Education

Institut de Science Financière et d'Assurances

Directeur : M. F. DE MARCHI

Directeur : M. le Professeur F. THEVENARD

Directeur : Mme C. FELIX

Directeur : M. Hassan HAMMOURI

Directeur : M. le Professeur S. AKKOUCHE

Directeur : M. le Professeur G. TOMANOV

Directeur : M. le Professeur H. BEN HADID

Directeur : M. le Professeur J-C PLENET

Directeur : M. Y. VANPOULLE

Directeur : M. B. GUIDERDONI

Directeur : M. le Professeur E.PERRIN

Directeur : M. G. PIGNAULT

Directeur : M. le Professeur C. VITON

Directeur : M. le Professeur A. MOUGNIOTTE

Directeur : M. N. LEBOISNE

to my great family

Abstract

Stroke and multiple sclerosis are two of the most destructive neurological diseases of the central nervous system. Stroke is the second most common cause of death and the major cause of disability worldwide whereas multiple sclerosis is the most common non-traumatic disabling neurological disease of adulthood.

Magnetic resonance imaging is an important tool to distinguish healthy from pathological brain tissue in diagnosis, monitoring disease evolution, and decision-making in personalized treatment of patients with stroke or multiple sclerosis.

Predicting disease evolution in patients with stroke or multiple sclerosis is a challenge for clinicians that are about to decide on an appropriate individual treatment. The etiology, pathophysiology, symptoms, and evolution of stroke and multiple sclerosis are highly different. Therefore, in this thesis, the statistical methods used for the study of the two neurological diseases are different.

The first aim was the identification of the tissue at risk of infarction in patients with stroke. For this purpose, the classification methods (including machine learning methods) have been used on voxel-based imaging data. The data measured at hospital admission is performed to predict the infarction risk at one month. Next, the performances of the classification methods in identifying the tissue at a high risk of infarction were compared.

The second aim was to cluster patients with multiple sclerosis using an unsupervised method based on individual clinical and imaging trajectories plotted over five 5 years.

Clusters of trajectories would help identifying patients who may have an important progression; thus, to treat them with more effective drugs irrespective of the clinical subtypes.

The third and final aim of this thesis was to develop a predictive model for individual evolution of patients with multiple sclerosis based on demographic, clinical, and imaging data taken at study onset. The heterogeneity of disease evolution in patients with multiple sclerosis is an important challenge for the clinicians who seek to predict the disease evolution and decide on an appropriate individual treatment. For this purpose, the latent class linear mixed model was used to predict disease evolution considering individual and unobserved subgroup' variability in multiple sclerosis.

Keywords: Stroke, Multiple Sclerosis, Classification Methods, Trajectory Clustering, Predictive Modeling, Latent Class Linear Mixed Model

Résumé

L'accident vasculaire cérébral et la sclérose en plaques figurent parmi les maladies neurologiques les plus destructrices du système nerveux central. L'accident vasculaire cérébral est la deuxième cause de décès et la principale cause de handicap chez l'adulte dans le monde alors que la sclérose en plaques est la maladie neurologique non traumatique la plus fréquente chez l'adulte jeune.

L'imagerie par résonance magnétique est un outil important pour distinguer le tissu cérébral sain de tissu pathologique à des fins de diagnostic, de suivi de la maladie, et de prise de décision pour un traitement personnalisé des patients atteints d'accident vasculaire cérébral ou de sclérose en plaques.

La prédiction de l'évolution individuelle de la maladie chez les patients atteints d'accident vasculaire cérébral ou de sclérose en plaques constitue un défi pour les cliniciens avant de donner un traitement individuel approprié. Cette prédiction est possible avec des approches statistiques appropriées basées sur des informations cliniques et d'imagerie. Toutefois, l'étiologie, la physiopathologie, les symptômes et l'évolution dans l'accident vasculaire cérébral et la sclérose en plaques sont très différents. Par conséquent, dans cette thèse, les méthodes statistiques utilisées pour ces deux maladies neurologiques sont différentes.

Le premier objectif était l'identification du tissu à risque d'infarctus chez les patients atteints d'accident vasculaire cérébral. Pour cet objectif, les méthodes de classification (dont les méthodes de machine learning) ont été utilisées sur des données d'imagerie

mesurées à l'admission pour prédire le risque d'infarctus à un mois. Les performances des méthodes de classification ont été ensuite comparées dans un contexte d'identification de tissu à haut risque d'infarctus à partir de données humaines codées voxel par voxel.

Le deuxième objectif était de regrouper les patients atteints de sclérose en plaques avec une méthode non supervisée basée sur des trajectoires individuelles cliniques et d'imagerie tracées sur cinq ans. Les groupes de trajectoires aideraient à identifier les patients menacés d'importantes progressions et donc à leur donner des médicaments plus efficaces.

Le troisième et dernier objectif de la thèse était de développer un modèle prédictif pour l'évolution du handicap individuel des patients atteints de sclérose en plaques sur la base de données démographiques, cliniques et d'imagerie obtenues à l'inclusion. L'hétérogénéité des évolutions du handicap chez les patients atteints de sclérose en plaques est un important défi pour les cliniciens qui cherchent à prévoir l'évolution individuelle du handicap. Le modèle mixte linéaire à classes latentes a été utilisé donc pour prendre en compte la variabilité individuelle et la variabilité inobservée entre sous-groupes de sclérose en plaques.

Mots de clé: Accident vasculaire cérébral, Sclérose en plaques, Méthodes de classification, Regroupement des données longitudinales, Modélisation prédictive, Modèle mixte à classes latentes

Acknowledgements

First, I wish to thank my supervisor, Delphine Maucort-Boulch, who gave me the opportunity to undertake this period of research and for all her help in the planning, execution and writing up of all work in this thesis. I benefitted greatly from her experience, expertise and encouragement which have been fundamental for the completion of this thesis.

I also want to thank my co-supervisor, Dominique Sappey-Marinier who has always been willing to share his unique experience and knowledge. He generously accepted to provide me the imaging data of multiple sclerosis patients, helped to interpret the results and correct my thesis. I also want to thank Gabriel Kocevar who used his effort and time to send me the clinical and imaging data of multiple sclerosis patients used in this dissertation. He always replied quickly about my questions related to the data and helped me for the correction of the last version of thesis.

I am also very grateful to the reviewers, Nural Bekiroglu, Marc Hommel, and Jacques Felblinger for accepting to participate to the jury of thesis defense, and for their precious and constructive comments which helped to enhance my thesis. I also want to thank Françoise Durand-Dubief and Lemlih Ouchchane for being part of my jury. It has been such a great honour to present and explain my work to you.

I am extremely thankful to Open Health Institute for funding my thesis. Their help not only provided financial support but also made me more self-assured.

My office companion Mathieu has always been very pleasant and vigilant support not only regarding scientific work also regarding his help on my French grammar mistakes.

Brice, Benjamin, Fabien, Amna and Florence have kindly helped in difficult coding issues and in teaching for Master students. Zoé, Laurant Re, Emmanuelle, Stéphanie B., Joris, Mad-Helenie, Manon, Saman, and Laurent Ro have been an essential part of good working environment and always guaranteed a friendly and inspiring atmosphere. I am very grateful to Pascal Roy and René Ecochard for their endless support during my scholarship and post-doctoral position research. I thank Stéphanie R. very much, Michèle, Marie-Thé, and Paola for their help in official paper processes and for familial approach. I am also very grateful to Jean Ivaz who helped me to correct the grammar and structure of all thesis!

Finally, I must thank my family, friends, and Anthony, who have been very understanding and patient during the thesis period with reduced presence from my side. I thank my parents, my little sis and brother who made me cry when laughing when I called them many times while stressed. Olivia, Tugce, Tugba, Kemal, Ali, Burç and Pierre, thank you for the great memories that we have shared together. I will always remember and miss these good days. Burcu, I thank you for listening me at least 2-3 hours many times and keeping me sane. Anthony, thank you for being always generous with endless encouragement and advice about my thesis and life. Life is better with you.

Contents

Abstract	iii
Resumé	v
List of Abbreviations	xvii
Introduction	1
1 State of the Art	5
1.1 Classification methods	7
1.1.1 Support Vector Machine	9
1.1.2 Artificial Neural Networks	11
1.1.3 Random Forest	14
1.1.4 Adaptative Boosting	16
1.1.5 Logistic Regression	18
1.1.6 Evaluation criteria for methods comparison	19
1.2 Clustering methods for longitudinal data	22
1.2.1 Single and joint trajectories	23
1.2.2 Distance between two trajectories	24
1.2.3 K-means for trajectories	28
1.3 Predictive modeling for longitudinal data	29
1.3.1 The linear mixed model	29
1.3.2 The latent class linear mixed model	34
1.4 Imaging	39
1.4.1 Conventional Imaging	41
1.4.2 Advanced Imaging	45
1.5 Overview of Stroke	47
1.5.1 Introduction	47
1.5.2 Epidemiology	48
1.5.3 Pathophysiology	49
1.5.4 Imaging of Stroke	51
1.5.5 Disease-modifying therapies	53

1.6	Overview of Multiple Sclerosis	55
1.6.1	Introduction	55
1.6.2	Epidemiology	56
1.6.3	Physiopathology	58
1.6.4	Evolution and severity of clinical disability	60
1.6.5	Imaging in MS	65
1.6.6	Diagnosis	72
1.6.7	Prognosis	73
1.6.8	Disease-modifying therapies	76
2	Comparison of classification methods in stroke	77
2.1	Introduction	78
3	Clustering of individual trajectories in MS	117
3.1	Introduction	118
3.2	Materials	120
3.2.1	Clinical data	120
3.2.2	Image Acquisition and Processing	120
3.3	Methods	121
3.4	Application of clustering methods	123
3.4.1	Statistical analyses	124
3.5	Results	126
3.5.1	Descriptive results	126
3.5.2	Cluster analysis results	134
3.6	Discussion	145
4	Modelling of individual disability evolution in MS	149
4.1	Introduction	150
5	Discussion	183
	Bibliography	216
	List of publications	217

List of Figures

- 1.1 A visualization of Support Vector Machine classification, Source: [140]. 10
- 1.2 A neural network with one hidden layer. 11
- 1.3 An artificial neuron. 12
- 1.4 A binary decision tree where x_i is all data, Z_t , K_m , and L_s are subsets of X_i where $t, m, s < i$ and v_l are the variables performed in classification where $l = 1, 2, 3, 4$ 14
- 1.5 A random forest performed with n trees with corresponding n samples. 15
- 1.6 Adaptative boosting algorithm. 16
- 1.7 Magnetization in a MRI. Direction of external magnetic field is in the head-foot direction in the MRI. 39
- 1.8 Effect of magnetic field and radio frequency on protons in the hydrogen molecule. 40
- 1.9 (a) When a 90-degree pulse is applied, the net magnetization vector of the protons is flipped from the longitudinal plane to the transverse plane. (b) When a 90-degree pulse is turned off, the net magnetization vector of the protons is flipped from the transverse plane to the longitudinal plane. 41
- 1.10 A brain imaging obtained respectively with T1-Weighted Imaging, T2-Weighted Imaging and FLAIR. 43
- 1.11 Contrast agent concentration change over time measured with T2WI in a brain voxel. 44
- 1.12 Figure on the left side shows the diffusion of the water in space (isotropic diffusion). On the right side, the figure shows the diffusion in brain (anisotropic diffusion). λ_1 : axial diffusivity, $(\lambda_2 + \lambda_3)/2$: radial diffusivity, $(\lambda_1 + \lambda_2 + \lambda_3)/3$: mean diffusivity. 46
- 1.13 Ischemic core and penumbra after stroke. While neurons in the ischemic core are considered irreversibly damaged, neurons in the penumbra are salvageable and are potential targets for therapeutic interventions [43]. 47
- 1.14 Global distribution of stroke mortality rates [107]. 49
- 1.15 A neuron in the Central Nervous System. 50
- 1.16 MRI of the brain. DWI (image at left side), and ADC map (image at right side), showing high signal intensity on the DWI and low signal intensity on the ADC parameter in a patient with ischemic stroke [139]. 52
- 1.17 T2, FLAIR and DWI imaging in ischemic stroke [81]. 52
- 1.18 DWI/PWI mismatch [114]. 53
- 1.19 MRI of the ischemic penumbra in a patient with stroke [53]. 53
- 1.20 The estimated number of people with MS in 2013 in each country. 57
- 1.21 Estimates of the regional prevalence of multiple sclerosis on 1 January 2003, per 100 000 inhabitants, standardised by age, among French farmers [230]. 57

1.22	Figure at the left side: A brain image at sagittal plane where the green part shows the Corpus Callosum. Figure at the right side: A brain image at coronal plane indicating the lateral ventricles and corpus callosum.	59
1.23	Change in disability score over time in the four MS clinical subtypes: Relapsing-Remitting (RR), Secondary Progressive (SP), Primary Progressive (PP), Clinically Isolated Syndrome (CIS).	60
1.24	Inflammation and axonal loss evolution during the transition from RR to SP. Relapse frequency decreases giving rise to axonal loss followed by permanent disability [229].	61
1.25	T2WI (A), FLAIR (B), and contrast-enhanced T1WI in a 30-year-old woman RRMS [74].	66
1.26	MS lesion in corpus callosum on T2WI (A), FLAIR (B) and FLAIR in sagittal plane (C) [74].	67
1.27	Images from a control subject (top row) and a patient (bottom row): DTI raw image (left column), color-coded FA map (middle column), and MD map (right column). FA takes the highest value (marked with red color) in the control subject whereas there is less red color in patients with MS, this indicates neurodegeneration. The lesions around CC are more visible on FA and MD maps [75].	70
1.28	Serial T2-FLAIR images (top row) and spin-echo MR (bottom row) recorded from baseline to one year in acute multiple sclerosis lesion. FLAIR images show the evolution of lesion size that increases during the first 20 days and decreases afterwards. The spectroscopic graphs show a decrease in NAA from baseline up to 20 days and then stabilization up to 1 year. An increase in Cho during the first weeks followed by a partial recovery is also observed whereas Cr stays relatively stable at all time points [179].	71
2.1	The AUC_{prc} values calculated separately for treated and untreated patients with tPA by each classification methods. The AUC_{prc} values performed for treated patients indicated adding “t” at the end of the method name on the x axis.	80
3.1	The horizontal axis shows the time in month and the vertical axis EDSS score. The dotted lines represent the individual EDSS trajectories and the solid lines the mean EDSS trajectories of each MS subtype.	128
3.2	The individual (dotted line) and mean (thick solid line) of GMV, WMV, LL, and FA trajectories observed in CIS, RR, PP, and SP patients.	131
3.3	Regression analysis between conventional and advanced imaging parameters. The correlation coefficient value was indicated on the top of each graph. The red line corresponds to the linear regression line.	132
3.4	Regression analyses between EDSS and imaging parameters.	133
3.5	Clustering results with graphical representation and corresponding table for single EDSS trajectories of all clinical subtypes. The individual trajectories are presented with black lines and the mean trajectories of the clusters are presented with colored thick solid lines. The proportions of individuals assigned to each cluster are given at the top of the graph.	136

3.6	Clustering results with graphical representations for respectively EDSS & LL trajectories (left side), and EDSS & FA trajectories (right side) with corresponding clustering tables for all clinical subtypes. The individual trajectories are represented with black lines and the mean trajectories are represented with colored thick solid lines. The proportions of individuals assigned to each cluster are given at the top of the graph.	138
3.7	Mean trajectories of the clusters obtained with single EDSS score trajectories in CIS, RR, and SP patients.	142
3.8	Clustering results of respectively EDSS score & LL (at left side) and EDSS score & FA (at right side) trajectories obtained with CIS, RR, and SP patients.	144

List of Tables

1.1	TP: True Positive, FP: False Positive, FN: False Negative, TN: True Negative, M+= Number of sick observations, M-= Number of healthy observations, T+= Number of observations with positive result, T-= Number of observations with negative result.	19
1.2	The Expanded Disability Status Scale (EDSS) [105].	64
1.3	Schumaker criteria for the diagnosis of MS [193].	73
1.4	2010 revised McDonald Diagnostic Criteria for MS.	74
3.1	Demographic and clinical characteristics of 75 patients measured at study onset, Median (range).	127
3.2	Conventional and advanced imaging parameters measured in 75 patients at study onset, Median (1st quartile - 3rd quartile) (GMV: Gray Matter Volume, WMV: White Matter Volume, LL: Lesion Load, FA: Fractional Anisotropy).	130
3.3	Calinski-Harabatz criterion results obtained with 2, 3, 4, 5 and 6 clusters performed on single and joint trajectories in all subtypes and for CIS, RR and SP patients.	134
3.4	Clinical and imaging variables according to the clusters obtained with single EDSS trajectories. Median (1st, 3rd quartile). the p-value is the result of Kruskal-Wallis test which was performed to compare the medians of variable values in the three clusters.	140

List of Abbreviations

AD Axial Diffusivity.

ADA Adaptative Boosting.

ADC Apparent Diffusion Coefficient.

AIC Akaike's Information Criterion.

ANN Artificial Neural Network.

AUC_{prc} Area Under the Precision-Recall Curve.

AUC_{roc} Area Under the receiver operating Curve.

BBB Blood-Brain Barrier.

BIC Bayesian Information Criterion.

CAIC Consistent Akaike's Information Criterion.

CART Classification and Regression Trees.

CBF Cerebral Blood Flow.

CBV Cerebral Blood Volume.

CC Corpus Callosum.

CH Calinski-Harabatz.

Cho Choline.

CIS Clinically Isolated Syndrome.

cMRI Conventional Magnetic Resonance Imaging.

CNS Central Nervous System.

CSF Cerebro-Spinal Fluid.

Cr Creatine.

DIS Dissemination In Space.

DIT Dissemination In Time.

DMT Disease Modifying Therapies.

DTI Diffusion Tensor Magnetic Resonance Imaging.

DWI Diffusion Weighted Magnetic Resonance Imaging.

EDSS Expanded Disability Status Scale.

EM Expectation-Maximization.

FA Fractional Anisotropy.

FS Functional Systems.

GA Glatiramer Acetate.

Glx Glutamate and Glutamine.

GMV Gray Matter Volume.

HLA Human Leukocyte Antigen.

IFN Interferon.

Lac Lactate.

LCMM Latent Class Linear Mixed Model.

LL Lesion Load.

LMM Linear Mixed Model.

LR Logistic Regression

MD Mean Diffusivity.

MHC Major Histocompatibility Complex.

MSFC Multiple Sclerosis Functional Composite.

ML Maximum Likelihood.

MR Magnetic Resonance.

MRI Magnetic Resonance Imaging.

MRS proton Magnetic Resonance Spectroscopy.

MRSI proton Magnetic Resonance Spectroscopic Imaging.

MS Multiple Sclerosis.

MTT Mean Transit Time.

NAA N-Acetyl-Aspartate.

NAWM Normal Appearing White Matter.

NIHSS National Institutes of Health Stroke Scale.

PASAT Paced Serial Addition Test.

PP Primary Progressive.

PWI Perfusion-Weighted Magnetic Resonance Imaging.

RD Radial Diffusivity.

RF Random Forest.

ROI Region Of Interest.

RR Relapsing Remitting.

SP Secondary Progressive.

SVM Support Vector Machines.

TNP True Negative Rate.

tPA Tissue Plasminogen Activator.

TPR True Positive Rate.

TTP Time To Peak of The Tissue Concentration-Time Curve.

T1WI T1-Weighted Imaging

T2WI T2-Weighted Imaging

T25FW Timed 25-Foot Walk.

UVR Ultraviolet Radiation.

VOI Volume Of Interest.

WMV White Matter Volume.

9HPT Nine-Hole Peg Test.

Introduction

The main objective is the application of appropriate statistical approaches in patients with neurological diseases such as ischemic stroke and multiple sclerosis to better characterize the disease evolution on the basis of demographic, clinical, and imaging data taken at study onset.

Ischemic stroke is defined as one of the most destructive neurological diseases and is the second most common cause of death and a major cause of disability worldwide [53]. Ischemic stroke occurs when the brain blood flow stops or is limited by a clot. The brain cells close to this area begin to die as they stop getting the oxygen and nutrients they need to function. One of the risk factors is ageing. Almost 75% of all strokes occur in people over the age of 65 and the risk of having a stroke more than doubles each decade after the age of 55. Additional risk factors for ischemic stroke are hypertension, diabetes, ischemic heart disease, atrial fibrillation, valvular heart disease, and cigarette smoking. Stroke has a negative impact on psychological and social situations such as problems with family relationships, deterioration in sex life, economic difficulties related to loss of the work, and deterioration in leisure activities [45].

Multiple sclerosis (MS) is the most frequent disabling neurological disease in young adults [119]. MS is a demyelinating inflammatory, chronic disease of the central nervous system. While its etiology remains unknown, it has been believed that the disease is triggered by environmental factors in genetically predisposed persons [209] [167]. There is no cure for multiple sclerosis. However, treatments can help speed recovery from clinical symptoms and modify the evolution of the disease. The evolution of disability is very different

from one patient to another [80]. Today's neurologists' challenge is to predict individual disability evolution on the basis of clinical, biological and imaging data.

Magnetic Resonance Imaging (MRI) is an important tool for diagnosis, monitoring the disease evolution, and decision-making in the personalized treatment of patients with stroke or MS. Various MRI parameters have been developed to obtain interesting information on a patient's state. In stroke, diffusion-weighted imaging (DWI) provides a marker of irreversibly damaged areas whereas perfusion-weighted imaging (PWI) is recommended to characterize the status of brain tissue blood supply [205] [240] [91]. In MS, the conventional imaging parameters (T1 and T2 Weighted Imaging, and FLAIR) have a high sensitivity in detecting lesioned area. Besides, the advanced imaging parameters (Diffusion Tensor Imaging and Magnetic Resonance Spectroscopic Imaging) are highly specific in detecting structural, functional, and metabolic changes that cannot be observed with conventional imaging [62] [56].

Chapter 1 details the statistical approaches such as the classification and clustering methods as well as the predictive models, respectively, in Section 1.1, Section 1.2, and Section 1.3. The section 1.4 provides the Magnetic Resonance Imaging and withal its sequences measured with conventional and advanced imaging. Afterward, brief reviews of stroke and multiple sclerosis -including their epidemiology, physiopathology, imaging, and disease-modifying therapies- are given in Sections 1.5 and 1.6.

The following three chapters (Chapter 2, Chapter 3, and Chapter 4) provide the three main contributions of the thesis :

- (a) Identification of the tissue at high risk of infarction in patients with ischemic stroke.

Identifying the tissue at high risk of infarction at one month may help the clinician treat stroke with more appropriate drugs. For this aim, we used five classification methods : four machine learning methods (Support Vector Machine, Artificial Neural Networks, Random Forest, and Adaptive Boosting) and one regression

model (Logistic Regression). The classification methods were performed based on voxel-based imaging data measured at hospital admission. The methods used to identify the tissue at high risk of infarction on voxel-based human data were compared mainly by the area under the precision-recall curve. The area under the ROC curve, sensibility, and specificity were also given to allow comparisons with previous results. Details on the classification methods used with stroke data, their results, and a discussion are given in Chapter 2.

- (b) Clustering patients with multiple sclerosis on the basis of longitudinal clinical and imaging data.

There is no cure for MS; it is a lifelong disease that lasts until death. Thus, to observe changes in disability and decide on treatment, patients with MS have to be constantly followed up from diagnosis until death. For this purpose, clustering methods were used with longitudinal multiple sclerosis data, not for a given time point. As unsupervised approach, an extension of the k-means method was used to cluster single clinical and joint clinical & imaging trajectories. The clusters may help identifying distinct clinical forms of evolution (e.g., stable, progressive, and regressive evolution). Such a classification is useful, for example, to identify the patients who may have an important progression and could benefit from early intervention with more effective drugs.

Details on the implementation of the clustering methods on data on patients with MS, the results, and a discussion are presented in Chapter 3.

- (c) Develop a statistical patient-specific model that predicts the progress of disability in multiple sclerosis on the basis of demographic, clinical, and multimodal imaging variables.

Clinicians need to predict short and long-term disability evolution in both the design of treatment trials with or without harmful therapies and to provide prognostic

advice to individual patients.

One of the difficulties is the heterogeneity of individual clinical evolutions. Therefore, predicting individual disability courses is extremely difficult. For this, we used the latent class linear mixed model in order to consider individual variability as well as unobserved subgroups variability in patients with multiple sclerosis.

Another difficulty of predictive modeling of disability evolution is the clinico-radiological paradox. Indeed, the changes in the imaging information measured with lesion load, white and gray matter volume, and diffusion measures are not directly reflected in the clinical disability. For example, an increase of lesion load is not always correlated with the clinical symptoms. However, many studies have found that some imaging variables were significantly related to clinical disability. For this reason, we established first a clinical model with clinical and demographic data. Second, the gray matter volume, the lesion load, and the fractional anisotropy were used to establish the imaging model. Finally, the imaging variables were added to the clinical model to investigate their additional contribution to model the evolution of individual disability .

The predictive models for disability evolution in multiple sclerosis are developed in Chapter 4. Details of the model implementation, the results and a discussion are also presented in the same chapter.

To conclude, we summarize the thesis by discussing the results in Chapter 5. The supplementary information and the figures are provided in the Appendix.

Chapter 1

State of the Art

Contents

1.1	Classification methods	7
1.1.1	Support Vector Machine	9
1.1.2	Artificial Neural Networks	11
1.1.3	Random Forest	14
1.1.4	Adaptative Boosting	16
1.1.5	Logistic Regression	18
1.1.6	Evaluation criteria for methods comparison	19
1.2	Clustering methods for longitudinal data	22
1.2.1	Single and joint trajectories	23
1.2.2	Distance between two trajectories	24
1.2.3	K-means for trajectories	28
1.3	Predictive modeling for longitudinal data	29
1.3.1	The linear mixed model	29
1.3.2	The latent class linear mixed model	34
1.4	Imaging	39
1.4.1	Conventional Imaging	41

1.4.2	Advanced Imaging	45
1.5	Overview of Stroke	47
1.5.1	Introduction	47
1.5.2	Epidemiology	48
1.5.3	Pathophysiology	49
1.5.4	Imaging of Stroke	51
1.5.5	Disease-modifying therapies	53
1.6	Overview of Multiple Sclerosis	55
1.6.1	Introduction	55
1.6.2	Epidemiology	56
1.6.3	Physiopathology	58
1.6.4	Evolution and severity of clinical disability	60
1.6.5	Imaging in MS	65
1.6.6	Diagnosis	72
1.6.7	Prognosis	73
1.6.8	Disease-modifying therapies	76

1.1 Classification methods

Classification methods, including machine learning methods, is a branch of Artificial Intelligence that focus on systems that learn from training data. Machine learning is used for many applications of pattern classification and recognition in health, computer science, and economics such as the identification of cancerous cells, prediction of financial indices as currency exchange rates, sexing of faces, and speech recognition. The global aim of machine learning methods is to compute, learn, remember, and optimize in the same way as a human brain.

Constructing a classifier requires two steps. First, the training set is used to decide how the classifiers or the observations ought to be weighted and combined in order to better separate the observations into various classes. Second, the weights determined in the training set are applied to a set of observations to predict their classes.

This section presents the classification methods that are appropriate for the binary outcome y_i (i.e. -1 vs 1 or healthy vs infarcted tissue). For a given observation set $\{(x_1, y_1), \dots, (x_i, y_i), \dots, (x_n, y_n)\}$, where x_i denotes the i th observation and y_i is the label that corresponds to the i th observation for $i \in 1, \dots, n$. The classification methods predict the outcome \hat{y}_i on the basis of m independent variables x_{im} measured for each observation i (See Equation 1.1). The training process continues until the method achieves a desired level of accuracy on the training data.

$$x_{im} = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iM} \end{pmatrix} = \begin{pmatrix} x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & \dots & x_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1M} & x_{2M} & \dots & x_{nM} \end{pmatrix} \quad (1.1)$$

This section provides an overview of four machine learning methods (Support Vector Machine, SVM; Artificial Neural Networks, ANN; Random Forest, RF; and Adaptive

Boosting, ADA) and one regression model, the Logistic Regression (LR). These four machine learning methods are frequently used for binary classification and LR is an easily applicable method, compared to machine learning algorithms, but that does not require parameter estimation before the fit of the method.

1.1.1 Support Vector Machine

SVM was developed by Cortes & Vapnik for binary classification [42]. SVM aims to find an optimal linear hyperplane to separate observations into two classes with the help of two classes' closest points (also called support vectors) (See Figure 1.1). The support vectors help choosing the best border either through maximizing the distance between the border and the support vectors or through minimizing the number of misclassified observations. This means that SVM requires the solution of the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}w^T w + C \sum_{i=1}^n \epsilon_i & (1.2) \\ \text{respecting} \quad & y_i(w^T \phi(x_i) + b) \geq 1 - \epsilon_i, \\ & \epsilon_i \geq 0. \end{aligned}$$

where

- w is the normal (orthogonal) vector to the border,
- $C > 0$ the regularization constant that controls the trade-off between the margin optimization and the number of observations incorrectly classified,
- ϵ_i is the misclassification rate (the number of the misclassified observations / the number of observations),
- ϕ is the projection function of the observations in the higher dimensional space.

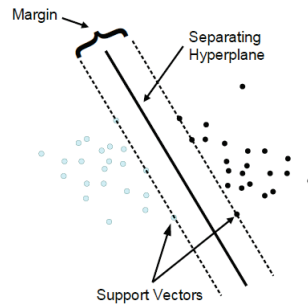


Figure 1.1: A visualization of Support Vector Machine classification, Source: [140].

When the observations cannot be separated into two dimensions by a linear hyperplane, they are projected in a higher dimensional space H where a linear separation is possible. The observations appear only inside dot products with other observations. More precisely, when a projection $\phi : X \rightarrow H$ is used, the dot product of two observations x_i and x_j is represented by a kernel function K :

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j) \quad (1.3)$$

There are four basic kernel functions: linear, polynomial, radial basis, and sigmoid functions. In this thesis, the radial basis function (RBF) was performed because it is the most used and recommended because of its good performance for complex data and the few number of parameters [234]. RBF is defined as below:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (1.4)$$

where γ is the kernel parameter.

The signed distance to the border provides the class of an observation. A zero distance corresponds to an observation located right on the border : a positive distance means it is located in one of the classes (for example, the sick side), whereas a negative distance means it is located in the other class (thus, healthy side). The border is defined as $\langle w, \phi(x) \rangle + b$, thus, the decision function [111] is:

$$f(x) = \text{sign}(\langle w, \phi(x) \rangle + b) \quad (1.5)$$

where b is a constant.

1.1.2 Artificial Neural Networks

ANN constitute a mathematical representation of natural neural networks composed of artificial neurons (See Figure 1.2) [101]. In the network, artificial neurons are organized in layers. Each network is composed of several types of layers: (1) an input layer composed of all inputs, (2) an output layer gives the final outcomes \hat{y}_i , and (3) one (or more) hidden layers between the input and the output layer consist(s) of a set of neurons that process the data and are connected to the input and output layers (See Figure 1.2). Artificial neurons are interconnected by weights -corresponding to the strengths between neurons- in such a way that information propagates from one layer to the next.

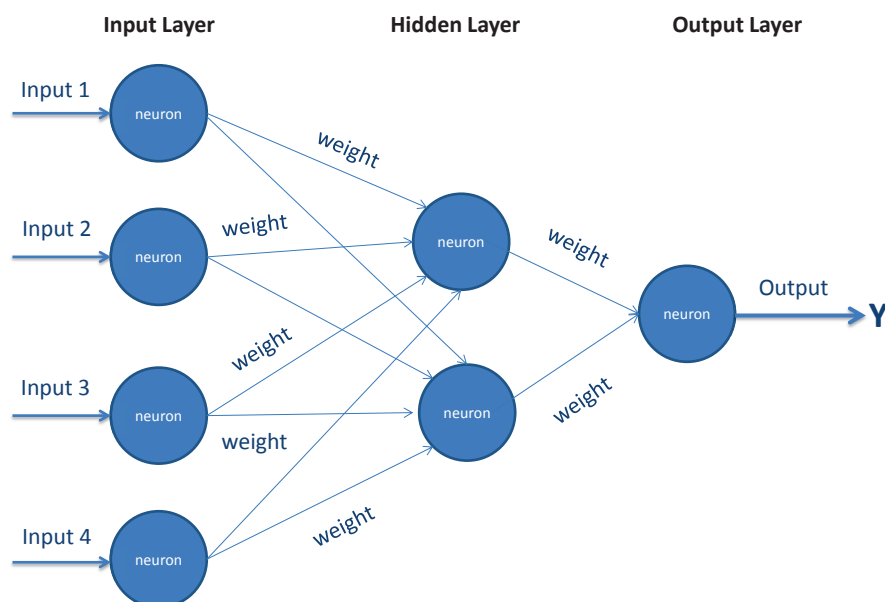


Figure 1.2: A neural network with one hidden layer.

A basic feature of ANN is a single artificial neuron (See Figure 1.3). An artificial neuron receives the observations x_i (representing dendrites in a natural neuron) and multiply them by the associated weight w_i that can vary during the learning process. The artificial neuron calculates afterward the weighted sum

$$\text{sum} = \sum_{i=1}^n x_i \times w_i \quad (1.6)$$

where x_i is the i th observation and w_i is the weight associated with the i th input.

Afterward, in a two-class classification problem, the weighted sum can be dichotomized with a threshold value b in the activation function. The neuron gives the output 1 if the weighted sum is above the threshold. Otherwise, the output is equal to 0.

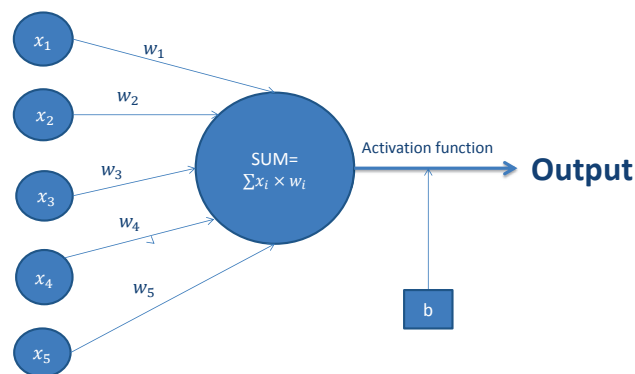


Figure 1.3: An artificial neuron.

The supervised learning algorithm can be performed on Artificial Neural Networks, thus ANN is able to learn from examples and adapt [82] [102]. The basic principle of learning from examples is to reduce the cost function. The cost function E is defined as half of the sum of the squared difference between the predicted outcome and the observed value :

$$E = \frac{1}{2} \times \sum_{i=1}^n (\hat{y}_i - d_i)^2 \quad (1.7)$$

where \hat{y}_i is the predicted output and d_i is the observed value of the observation i .

ANN improve their performances by providing a better "relationship" between input and output, which means providing better weights between neurons [82] [102]. A "back-propagation" learning strategy [90] was realized by updating the weights according to the error obtained by the network at the output. The back-propagation learning strategy is defined as :

1. the weights and threshold are initialized with small random values
2. network computes the outcome \hat{y}_i
3. the weights are updated according to

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij} \quad (1.8)$$

where w_{ij} is the weight between neuron i and neuron j in two different layers, t is the iteration number, and Δw_{ij} is a constant.

The commonly used method to find Δw_{ij} is gradient descent which aims to minimize the cost function E . Thus, Δw_{ij} is defined as :

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} \quad (1.9)$$

where η is the learning rate defined at the beginning of network.

The learning process continues until the minimum gradient of the error function is reached.

1.1.3 Random Forest

RF is an ensemble model which combines several decision trees to produce a more accurate prediction than a simple decision tree [23]. In this study, the binary classification trees (Classification and Regression Trees (CART)) [24] will be presented and used to classify the observations into two groups (See Figure 1.4).

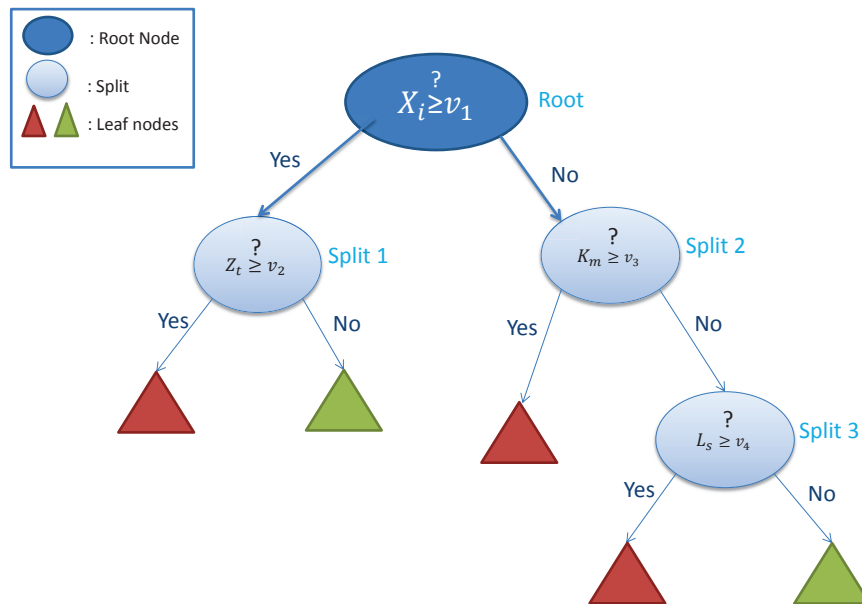


Figure 1.4: A binary decision tree where x_i is all data, Z_t , K_m , and L_s are subsets of X_i where $t, m, s < i$ and v_l are the variables performed in classification where $l = 1, 2, 3, 4$.

The aim of a binary classification tree is to separate the observations into two classes with the variables v_l at root node and splits. A subsample of variables is randomly created for each tree and its size is fixed a priori (mtry). The best variable is chosen within a subsample of variables for each split and root node according to the Gini Index. The decision tree optimizes the homogeneity at split and root node by minimizing the Gini Index defined as :

$$\text{Gini Index} = p_1(1 - p_1) + p_2(1 - p_2) \quad (1.10)$$

where p_1 and p_2 are respectively the proportions of observations in class 1 and class 2. Whenever one split becomes sufficiently homogeneous, the classification stops in this split.

RF builds several bootstrap samples from the original data and fits a decision tree to each sample (See Figure 1.5) [23]. The tree fitted for each sample is different from another tree due to the randomly chosen subsample of variables v_l . This makes RF more flexible and powerful. For a given observation, the outcome of the random forest \hat{y} is the majority vote of binary decision trees results.

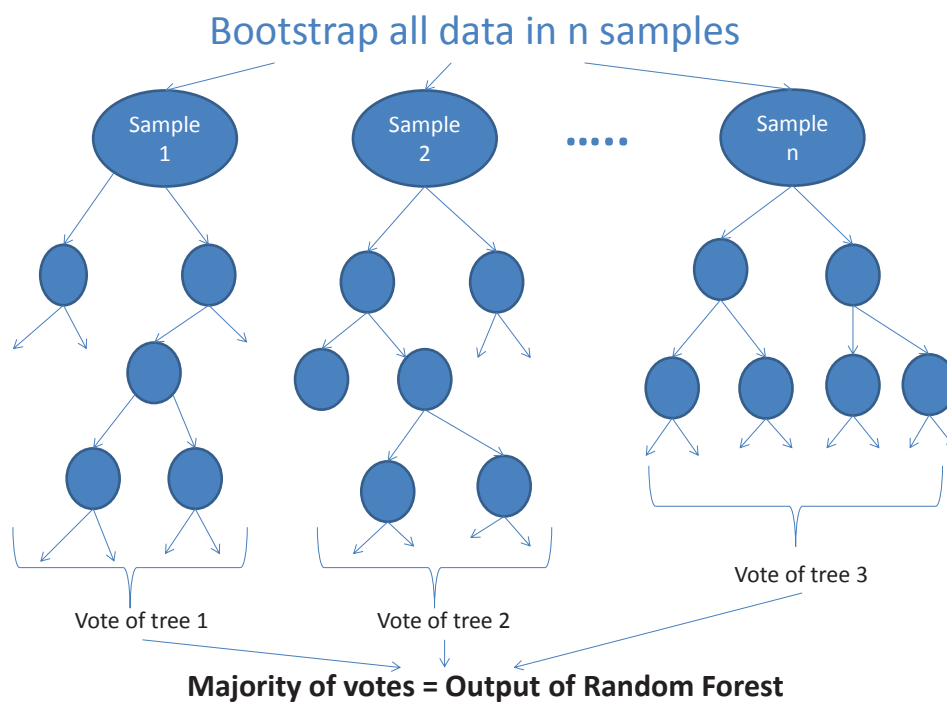


Figure 1.5: A random forest performed with n trees with corresponding n samples.

1.1.4 Adaptive Boosting

ADA method includes a boosting algorithm that refers to a family of algorithms which converts "weak" classifiers to "strong" classifiers by combining the predictions of different weak classifiers (See Figure 1.6) [191]. Thus, the prediction is much more accurate than any prediction by the weak classifiers.

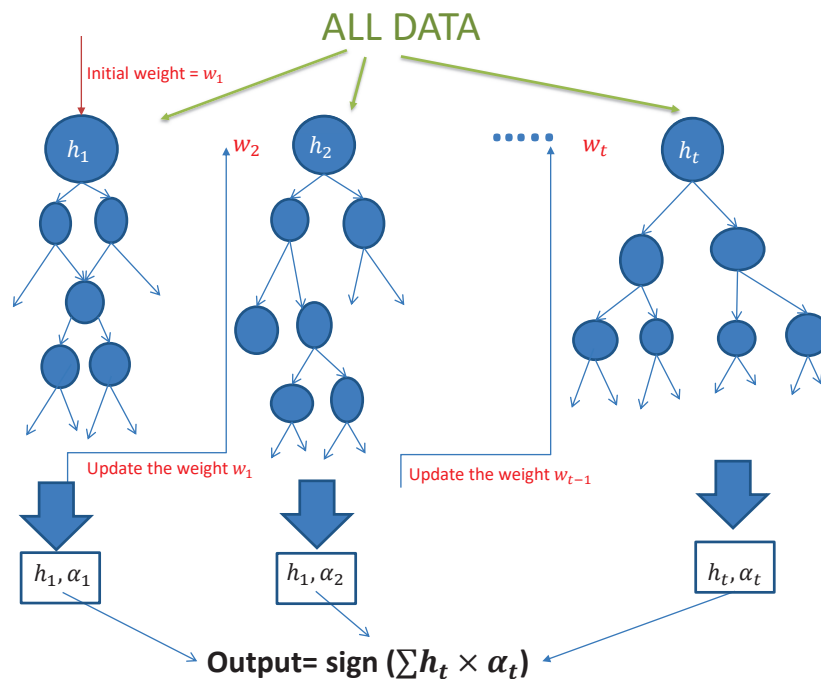


Figure 1.6: Adaptive boosting algorithm.

The ADA uses a set of decision trees h_t (called the weak classifiers) for $t \in \{1, \dots, T\}$ and corrects sequentially the classification weighting the misclassified observations obtained from the previous decision tree (See Figure 1.6). The classification of the first tree is performed with the initial weights which are the same for all observations. However, the weights of misclassified observations are increased after each decision tree classification to focus on the observations that are difficult to classify (See Equation 1.12). The performance of each tree is estimated according to the misclassification error of the decision tree ($\epsilon_t = Pr[h_t(x_i) \neq y_i]$) and the weights which are given to the observations (See Equation

1.11). For a given observation, the output is the result of the classification of each tree weighted by its performance (See Equation 1.13). The ADA algorithm [191] is defined as follows:

- Given $(x_i, y_i) = \{(x_1, y_1), \dots, (x_n, y_n)\}$ where $x_i \in X$ and $y_i \in \{-1, +1\}$
- Initialize $w_1(i) = 1/n$

For $t = 1, \dots, T$:

1. Train weak classifier (CART) using distribution w_t
2. Get weak hypothesis $h_t : X \rightarrow \{-1, +1\}$
3. Estimate

$$\alpha_t = \frac{1}{2} \ln\left(\frac{1 - \epsilon_t}{\epsilon_t}\right) \quad (1.11)$$

4. Update the weight

$$w_{t+1}(i) = \frac{w_t(i)}{Z_t} \times K_t = \frac{w_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t} \quad (1.12)$$

where Z_t is a normalization factor chosen so that w_{t+1} will be a distribution and $K_t = e^{-\alpha_t}$ if $h_t(x_i) = y_i$, $K_t = e^{+\alpha_t}$ if not.

- The output of algorithm is

$$\hat{y}_i = \text{sign}\left(\sum_{t=1}^T \alpha_t \times h_t(x)\right) \quad (1.13)$$

1.1.5 Logistic Regression

LR is an appropriate and easy-to-apply regression analysis when the binary output y_i is modeled by one or more than one independent variable(s) x_{im} . In LR, the probability of being in one class π_i is defined as:

$$\pi_i = \frac{1}{1 + e^{-(\beta_0 + \sum_{m=1}^M \beta_m \times x_{im})}} \quad (1.14)$$

where β_0 is the intercept and β_m the effect size of one unit change of the independent variable x_{im} on the output. When β_m is positive, it can be interpreted as a positive association. However, a negative β_m means that this variable is a negative association. $\beta_m = 0$ means that the independent variable m has no effect on the output.

The probability of being in one class π_i can be dichotomized then with a threshold value in order to obtain a binary output.

The parameter estimation for β_m can be obtained by maximizing the likelihood function is defined as:

$$L(\beta_m) = \prod_{i=1}^n \pi(x_{im})^{y_i} \times (1 - \pi(x_{im}))^{1-y_i} \quad (1.15)$$

LR assumes before the application of the method such as:

1. independence of observations. This means that the observations should not come from repeated measurements or matched data,
2. uncorrelated (independent) error terms,
3. no multicollinearity among independent variables (for example, weight and body mass index are correlated, so we can not use both in the same model).

1.1.6 Evaluation criteria for methods comparison

Let y_i be a binary observation where $y_i \in \{0, 1\}$ (e.g. $y_i = 1$ for sick observations and $y_i = 0$ for healthy observations). The classification methods provide the probability of being in one class (π_i) which can be dichotomized afterward with a threshold value τ . The observation is labeled 1 when the probability of being in one class is above the threshold of the biomarker τ , and otherwise the observation is labeled 0.

Table 1.1 is 2×2 contingency table and shows the observations with the classification method results.

	Sick Observation $y_i=1$	Healthy Observation $y_i=0$	
Positive result: ($\pi_i \geq \tau$)	TP	FP	T+=TP+FP
Negative result: ($\pi_i < \tau$)	FN	TN	T-=FN+TN
	M+=TP+FN	M-=FP+TN	

Table 1.1: TP: True Positive, FP: False Positive, FN: False Negative, TN: True Negative, M+= Number of sick observations, M-= Number of healthy observations, T+= Number of observations with positive result, T-= Number of observations with negative result.

Sensitivity is the ability of a classification method to classify correctly "sick" observations. Thus, sensitivity is the probability of having a positive result for a sick observation and is defined as :

$$Se = P(T+ | M+) = \frac{P(T+ \cap M+)}{P(M+)} = \frac{TP}{TP + FN} \quad (1.16)$$

Specificity is the ability of a classification method to classify correctly the "healthy" observations. Thus, specificity is the probability of having a negative result for a healthy observation and is defined as:

$$Sp = P(T- | M-) = \frac{P(T- \cap M-)}{P(M-)} = \frac{TN}{TN + FP} \quad (1.17)$$

Sensitivity and specificity depend on the threshold value τ . So, increasing the threshold value decreases the sensitivity and increases the specificity [213] [84].

The area under the Receiver Operating Curve (AUC_{roc}) plots Sensitivity as a function of 1-Specificity at different threshold points of a biomarker. Sensitivity and specificity are calculated for each threshold and the sensitivity/specificity pair creates a curve called ROC. The area under the ROC is a measure of how well a biomarker is able to distinguish between two binary responses.

$$AUC_{roc} = \int_{-\infty}^{+\infty} Se(\tau) d\{1 - Sp(\tau)\} \quad (1.18)$$

AUC_{roc} is an evaluation criterion which is independent of the imbalance in the data. AUC_{roc} can overestimate the performance of the method, for example, for a data where there are more healthy observations than sick observations. This problem can be overcome with the area under the Precision Recall Curve and some authors have recommended the use of the AUC_{pr} in addition of the AUC_{roc} in studies with low prevalence values [48] [155].

The area under the Precision Recall Curve (AUC_{pr}) plots Precision as a function of Sensitivity at different threshold points of a biomarker whenever Precision is defined as:

$$\text{Precision} = P(M+ | T+) = \frac{P(M+ \cap T+)}{P(T+)} = \frac{TP}{TP + FP} \quad (1.19)$$

The area under Precision Recall Curve (AUC_{pr}) is defined as:

$$AUC_{pr} = \int_{-\infty}^{+\infty} \text{Precision}(\tau) d\{Se(\tau)\} \quad (1.20)$$

Unlike sensitivity and specificity, AUC_{roc} and AUC_{pr} do not depend on the threshold value.

1.2 Clustering methods for longitudinal data

In data analysis, clustering is a standard procedure to explore the characteristics of groups of individuals. The aim is to cluster the more similar individuals in the same group and the more dissimilar individuals in different clusters. The clustering analysis can be categorized into two main groups: non-parametric and model-based. In this thesis, we will present first non-parametric clustering methods. Second, model-based clustering presented in the following section.

Clustering methods are often confused with classification methods. In classification, the classes to which we want to assign the individuals are already predefined whereas, in clustering, the clusters are established during the process.

Clustering analyses can be performed on cross-sectional data as well as the longitudinal data. In health, social, or behavioral sciences, longitudinal studies are important to observe changes over time. Here, we clustered the individuals according to their individual trajectories established with longitudinal data.

This section provides an overview of clustering algorithms for single and joint longitudinal data. We present first the single and joint trajectories in the section [1.2.1](#). Section [1.2.2](#) introduces the concept of distance between trajectories, data standardization, and selection of the number of clusters. Finally, Section [1.2.3](#) describes the k-means algorithm for longitudinal data.

1.2.1 Single and joint trajectories

Single trajectories

Let S be a set of n trajectories and y_t the longitudinal outcome measured at different time points t . A single trajectory (or single-variable trajectory) for a given individual i is defined as $y_i = (y_{i1}, y_{i2}, \dots, y_{it})$. For a given variable X , a single variable trajectory can be written as $y_{i.X} = (y_{i1X}, y_{i2X}, \dots, y_{itX})$.

Joint trajectories

A joint trajectory is a multi-variable trajectory obtained with at least 2 variables. Let $y_{i.A}, y_{i.B}, \dots, y_{i.M}$ be single trajectories for variable A, B, \dots, M of individual i . The joint trajectory of the individual i is defined as :

$$y_{i..} = \begin{pmatrix} y_{i.A} \\ y_{i.B} \\ \vdots \\ y_{i.M} \end{pmatrix} = \begin{pmatrix} y_{i1A} & y_{i2A} & \dots & y_{itA} \\ y_{i1B} & y_{i2B} & \dots & y_{itB} \\ \vdots & \vdots & \ddots & \vdots \\ y_{i1M} & y_{i2M} & \dots & y_{itM} \end{pmatrix} \quad (1.21)$$

where each line corresponds to a single trajectory for a variable and t represents time.

1.2.2 Distance between two trajectories

Several distance metrics are possible to determine the distance between two single or joint trajectories. However, we will present only two distance metrics frequently used in the literature: the Euclidian distance and the Manhattan distance.

Distance between two single trajectories

The Euclidian distance between two single trajectories y_i and y_j of individuals i and j is defined as:

$$Dist(y_i, y_j) = \sqrt{\frac{1}{t} \sum_{k=1}^t (y_{ik} - y_{jk})^2} \quad (1.22)$$

where t represents time.

The Manhattan distance is more robust than the Euclidian distance in case of unusual values like outliers [113] and is defined as:

$$Dist(y_i, y_j) = \sum_{k=1}^t |y_{ik} - y_{jk}| \quad (1.23)$$

Distance between two joint trajectories

Let $y_{1..}$ and $y_{2..}$ be the joint trajectories of individuals $i = 1$ and $j = 2$. The distance between two joint trajectories is defined as $d(y_{i..}, y_{j..}) = d(y_{1..}, y_{2..})$ and can be considered as the distance between two matrices.

$$\begin{aligned}
d(y_{1..}, y_{2..}) &= \text{Dist} \left(\begin{bmatrix} y_{11A} & y_{12A} & \cdots & y_{1tA} \\ y_{11B} & y_{12B} & \cdots & y_{1tB} \\ \vdots & \vdots & \ddots & \vdots \\ y_{11M} & y_{12M} & \cdots & y_{1tM} \end{bmatrix}, \begin{bmatrix} y_{21A} & y_{22A} & \cdots & y_{2tA} \\ y_{21B} & y_{22B} & \cdots & y_{2tB} \\ \vdots & \vdots & \ddots & \vdots \\ y_{21M} & y_{22M} & \cdots & y_{2tM} \end{bmatrix} \right) \\
&= \text{Dist} (d_1.(y_{11.}, y_{21.}), d_2.(y_{12.}, y_{22.}), \dots, d_t.(y_{1t.}, y_{2t.}))
\end{aligned} \tag{1.24}$$

We compute first t distances between M rows of each column where M is the number of variables. So, t distances $d_t.(y_{1t.}, y_{2t.}) = (d_1.(y_{11.}, y_{21.}), d_2.(y_{12.}, y_{22.}), \dots, d_t.(y_{1t.}, y_{2t.}))$ are obtained. Then, these t distances are combined using a distance function $d(y_{1..}, y_{2..}) = \text{Dist}(d_1.(y_{11.}, y_{21.}), d_2.(y_{12.}, y_{22.}), \dots, d_t.(y_{1t.}, y_{2t.}))$.

The Euclidean distance between two joint trajectories $y_{i..}$ and $y_{j..}$ can be computed as:

$$\text{Dist}(y_{1..}, y_{2..}) = \sqrt{\sum_t \sum_X (y_{1tX} - y_{2tX})^2} \tag{1.25}$$

where t and X represent, respectively, the time and variable.

The Manhattan distance between two joint trajectories is defined as:

$$\text{Dist}(y_i, y_j) = \sum_t \sum_X |y_{itX} - y_{jtX}| \tag{1.26}$$

Data standardization

The variables that generate the joint trajectories are not always expressed in the same unit. The variable with the larger scale may have a heavier weight than the other variables. This is not desired in cluster analysis [58].

The solution is to normalize the single variable trajectories with the mean $\overline{y_{..X}}$ and the standard deviation $sd_{..X}$ of the variables calculated over all individuals and time points, not on a given individual or time point. The normalized single variable trajectory for variable X is defined as below:

$$y'_{itX} = \frac{y_{itX} - \overline{y_{..X}}}{sd_{..X}} \quad (1.27)$$

The normalized joint trajectory is defined as the matrix obtained with normalized single variable trajectories as shown below:

$$y'_{i..} = \begin{pmatrix} y'_{i.A} \\ y'_{i.B} \\ \vdots \\ y'_{i.M} \end{pmatrix} \quad (1.28)$$

Optimal number of clusters

The number of clusters should be set before the clustering is performed. The most frequently used criterion to determine the "optimal" number of clusters are the Calinski & Harabatsz criterion $C(g)$ for a given cluster number g . Other versions have been published such as Kryszczuk $C_K(g)$ and Genolini variants $C_G(g)$ [29] [117] [78] [144] [197]. In addition to Calinski & Harabatsz criterion and its variants, Davies & Boudin [47] and Ray & Turi [170] criteria may be also used to evaluate the discrimination ability for the purpose of choosing the number of clusters.

The Calinski & Harabatsz criterion combines the within (**W**) and between (**B**) variance matrix defined as:

$$\begin{aligned} (\mathbf{W}) &= \sum_{g=1}^G \sum_{i=1}^{n_g} (\bar{y}_{gi} - \bar{y}_g) \times (\bar{y}_{gi} - y_g)^T \\ (\mathbf{B}) &= \sum_{g=1}^G n_g (\bar{y}_g - \bar{y}) \times (\bar{y}_g - \bar{y})^T \end{aligned} \tag{1.29}$$

where g represents the cluster in $\{1, \dots, G\}$ and n_g is the number of trajectories in the cluster g .

So, the Calinski & Harabatsz criteria are defined as

$$\begin{aligned} C(g) &= \frac{\text{Trace}(B)}{\text{Trace}(W)} \times \frac{n-g}{g-1} \\ C_K(g) &= \frac{\text{Trace}(B)}{\text{Trace}(W)} \times \frac{n-1}{n-g} \\ C_G(g) &= \frac{\text{Trace}(B)}{\text{Trace}(W)} \times \frac{n-g}{\sqrt{g-1}} \end{aligned} \tag{1.30}$$

where $\text{Trace}(B)$ and $\text{Trace}(W)$ represent, respectively, the sum of the diagonal coefficients of B and W , n is the total number of trajectories and g represents the number of clusters.

A high value of $\text{Trace}(B)$ indicates a great distance between clusters while a low value of $\text{Trace}(W)$ indicates close trajectories in each cluster. Thus, a high value of $\text{Trace}(B)$ and a low value of $\text{Trace}(W)$ mean that the number of cluster g ensures good discrimination. The cluster number k is chosen so as to maximize the $C(g), C_K(g)$, and $C_G(g)$.

1.2.3 K-means for trajectories

K-means is an unsupervised learning algorithm that classifies a given data set into a certain number of clusters fixed a priori [133]. K-means belongs to the Expectation-Maximization (EM) class and follows two steps (E-step and M-step) as below:

1. (E-step) A trajectory that corresponds to an individual is randomly assigned to a cluster and the mean of the trajectories is defined as the center of the cluster
2. (M-step) The distance between a trajectory and the center of the clusters is calculated and then each trajectory is assigned to the nearest cluster.

The EM algorithm continues until the trajectories stop to change place.

K-means algorithm can be used to cluster single and joint trajectories using software R-project (<https://www.r-project.org/>) (respectively, with the **kml** and **kml3d** packages [78]).

1.3 Predictive modeling for longitudinal data

In this section, we will present the linear mixed model and its extension called the latent class linear mixed model. These models are useful for longitudinal data and allow for individual differences and the correlation due to repeated measures. The advantage of the latent class linear mixed model versus the standard linear mixed model is that both the fixed effects and the distribution of random effects can be class-specific. The class-specific changes may be due to unobserved groups called latent classes. So, there are two types of heterogeneity in the latent class linear mixed models: heterogeneity between individuals and heterogeneity between latent classes.

We will present first the linear mixed model and the way parameter estimation and individual prediction are performed. The transition to the latent class linear mixed model will be more straightforward. Second, we will present the latent class mixed model, parameter estimation, the choice of the number of latent class, the posterior probability of being in the latent classes, the marginal and subject-specific predictions, and the model evolution criteria.

1.3.1 The linear mixed model

The linear mixed model [120] is a flexible approach to take into account correlated data due to repeated measures obtained at successive time points. The linear mixed model includes jointly inter-individual changes (random effects) apart from the population effect which is the same for all individuals (fixed effects). The random effect refers to the way individual measurements spread out from the common effects (with different intercepts or/and slope for each individual).

Let Y_i be the repeated outcome for individual i as $Y_i = (Y_{i1}, \dots, Y_{in_i})$ where n_i represents the number of the repeated measures. The linear mixed model for subject i can be written as:

$$Y_i = X_i\beta + Z_iu_i + \epsilon_i \quad (1.31)$$

where X_i is a $n_i \times p$ corresponds to the matrix for the vector of fixed effects β in length p and Z_i is a $n_i \times q$ design matrix for the vector of random effects u_i in length q . The linear mixed models assume that the population is homogeneous and that the random effects are normally distributed as $u_i \sim N(0, B)$. ϵ_i are the measurement errors that are assumed to be normally distributed with mean zero and covariance matrix $\sigma^2 I_{n_i}$ (n_i is the number of repeated measures of individual i) and to be independent from the random effects u_i .

Parameter Estimation

Parameter estimation in a linear mixed model is possible with the maximum likelihood (ML) method. First, we define the marginal distribution of Y_i :

$$Y_i \sim N(X_i\beta, V_i = Z_iBZ_i^T + \sigma^2I_i) \quad (1.32)$$

This notation is used for estimating a set of fixed effect parameters θ (β , B and σ^2). The likelihood function for the set of parameters θ is characterized as below:

$$\begin{aligned} L_i(\theta) &= \phi_i(Y_i; \theta) \\ &= \prod_{i=1}^N \left(\frac{1}{2\pi} \right)^{\frac{n_i}{2}} \times |V_i|^{-1/2} \times \exp\left(-\frac{1}{2}(Y_i - X_i\beta)^T \times V_i^{-1} \times (Y_i - X_i\beta)\right) \end{aligned} \quad (1.33)$$

where ϕ_i is the density function at individual level with individual specific mean and variance $\mu_i = X_i\beta$ and $V_i = Z_iBZ_i^T + \sigma^2I_i$ and N is the number of individuals.

The parameters can be estimated maximizing log-likelihood $\log(L_i(\theta))$. The maximization of the log-likelihood requires an iterative algorithm. Two possible algorithms may be used: the Expectation - Maximization (EM) algorithm and the Newton-Raphson algorithm [51]

[68]. The latter is generally preferred because of its convergence faster [126]. For this reason, the Marquardt algorithm, a Newton-Raphson-like method was used in this work.

-Marquardt algorithm

Marquardt algorithm [134] is an iterative method of the Newton-Raphson family that is used to find the stationary points (local maximum or minimum) of a differential function f . To find the stationary points x which provide $f'(x) = 0$, we need to calculate the second derivative of the twice-differentiable function f . At the stationary points, the gradient is equal to zero as $-\frac{\partial f(x)}{\partial x} = -\nabla(x)$.

Here, function f was defined as the log-likelihood function and we searched the convergence through

1. the gradient (first derivative) of the log-likelihood function for the θ values as

$$-\frac{\partial \log(L_i(\theta))}{\partial \theta} = -\nabla(\theta) \quad (1.34)$$

2. the second derivative of the log-likelihood function (also called Hessian)

$$-\frac{\partial^2 \log(L_i(\theta))}{\partial \theta^2} = H(\theta) \quad (1.35)$$

The Marquardt algorithm which is defined in the equation 1.36 updates the parameters θ until the convergence of the function.

$$\theta^{(k+1)} = \theta^{(k)} - \delta \times (\tilde{H}^{-1}(\theta^k)) \times \nabla(\theta^{(k)}) \quad (1.36)$$

where δ is normally equal to 1 but may be modified to ensure the log-likelihood is improved at each iteration and \tilde{H}^{-1} is the inverse of the positive defined Hessian matrix at diagonal terms where for $\tilde{H} = \tilde{H}_{ij}$, $\tilde{H}_{ii} = H_{ii} + \lambda \times [(1 - \eta)|H_{ii}| + \eta \times \text{tr}(H)]$.

Initial values of λ and η are fixed as 0.01 by default, they may be decreased when \tilde{H} is chosen positive and increased if not.

Convergence is reached when all three following criteria are satisfied (1) $\sum_{j=1}^m (\theta_j^{(k)} - \theta_j^{(k-1)})^2 \leq 10^{-4}$, (2) $|L^{(k)} - L^{(k-1)}| \leq 10^{-4}$, and (3) $\nabla(\theta^{(k)})^T (H^{(k)})^{-1} \nabla(\theta^{(k)}) \leq 10^{-5}$.

Prediction

First, we will present the prediction of the random effects by the Empirical Bayes method, then, the marginal and individual (subject-specific) predictions.

Empirical Bayes estimates of random effects:

The marginal distribution of Y and the distribution of the random effect are respectively defined as $Y_i \sim N(X_i\beta, V_i)$ and $u_i \sim N(0, B)$. So, the covariance between Y_i and u_i can be written as follows:

$$\begin{aligned}
 cov(Y_i, u_i) &= cov(X_i\beta + Z_i u_i + \epsilon_i, u_i) \\
 &= cov(X_i\beta, u_i) + Z_i var(u_i, u_i) + cov(\epsilon_i, u_i) \\
 &= 0 + Z_i B + 0 \\
 &= Z_i B
 \end{aligned} \tag{1.37}$$

The covariance between Y_i and u_i is used afterward to obtain the joint distribution of Y_i and u_i :

$$\begin{pmatrix} Y_i \\ u_i \end{pmatrix} \sim N \left(\begin{pmatrix} X_i\beta \\ 0 \end{pmatrix}, \begin{pmatrix} V_i & Z_i B \\ B Z_i^T & B \end{pmatrix} \right) \tag{1.38}$$

This distribution helps obtaining the esperance of random effects conditional to the response variable as defined below:

$$\begin{aligned}
E(u_i, Y_i) &= E(u_i) + \text{cov}(u_i, Y_i) \text{var}(Y_i)^{-1} [Y_i - E(Y_i)] \\
&= 0 + BZ_i^T V_i^{-1} (Y_i - X_i \beta) \\
&= BZ_i^T V_i^{-1} (Y_i - X_i \beta)
\end{aligned} \tag{1.39}$$

So, the Empirical Bayes estimates of the random effects u_i can be written including the estimated fixed effect parameter $\hat{\beta}$ as

$$\hat{u}_i = BZ_i^T V_i^{-1} (Y_i - X_i \hat{\beta}) \tag{1.40}$$

Marginal and subject-specific predictions:

The marginal and subject-specific predictions can be calculated by replacing the fixed parameters, and Bayes estimates of the random effects by their estimated values ($\hat{\beta}$ and \hat{u}_i) in the equation 1.31.

Marginal and subject-specific predictions in the linear mixed model are calculated as

$$\begin{aligned}
\hat{Y}_{ij}^{(M)} &= X_{ij}^T \hat{\beta} \\
\hat{Y}_{ij}^{(SS)} &= X_{ij}^T \hat{\beta} + Z_{ij}^T \hat{u}_i
\end{aligned} \tag{1.41}$$

where i indicates the individual and j the time.

So, the marginal and subject-specific residuals are defined as:

$$\begin{aligned}
\hat{R}_{ij}^{(M)} &= Y_{ij} - \hat{Y}_{ij}^{(M)} \\
\hat{R}_{ij}^{(SS)} &= Y_{ij} - \hat{Y}_{ij}^{(SS)}
\end{aligned} \tag{1.42}$$

1.3.2 The latent class linear mixed model

Linear mixed models assume that there is a unique distribution of the population with a unique mean for all individuals. This means that the population is homogeneous among all individuals. However, Verbeke and Lesaffre [227] introduced the mixture of random effects that corresponds to the unobserved variability between sub-groups which may be present in the given data. The latent class linear mixed model is an extension of the linear mixed models for a heterogeneous population divided into G unobserved subgroups (also called latent classes). So, the latent class linear mixed model provides g different distributions, so g different means for each subgroup $g \in \{1, \dots, G\}$ within the population with a common covariance matrix B .

The latent class linear mixed model can be written as

$$Y_{ij}|c_i=g = X_{1ij}\beta + X_{2ij}\gamma_g + Z_{ij}u_{ig} + \epsilon_{ij} \quad (1.43)$$

where X_{1ij} and X_{2ij} are the covariate matrices and associated with the vector of fixed effects, at population and latent class levels, respectively. Z_{ij} is the class-specific covariate matrix of class g and is associated with the random effects $u_{ig} \sim N(\mu_g, B_g)$ where u_g and B_g are, respectively, the class-specific mean and the covariance matrix.

The probability of belonging to latent class g can be calculated with the multinomial logistic model:

$$p_{ig} = P(c_i = g|X_{ci}) = \frac{e^{\xi_{0g} + X_{ci}^T \xi_{1g}}}{\sum_{k=1}^G e^{\xi_{0k} + X_{ci}^T \xi_{1k}}} \quad (1.44)$$

where c_i represents the latent class for individual i , ξ_{0g} is the intercept of the latent class g , X_{ci} represents the individual class-specific covariate matrix, ξ_{1g} is a class specific parameter and G is the total number of latent classes. ξ_{0g} and ξ_{1g} are equal to 0 for identifiability [166].

Parameter Estimation

The difference between parameter estimation in the latent class linear mixed model versus the linear mixed model is that the parameters are estimated for each latent class.

In the latent class linear mixed model, the parameter estimation is possible with the maximum likelihood method [227] [149] [165] [116]. The likelihood function for an individual i for the set of parameters θ_G (β and γ_g) is defined as:

$$\begin{aligned} L_i(\theta_G) &= \sum_{g=1}^G p_{ig} \times \phi_{ig}(Y_i | c_i = g; \theta_G) \\ &= \sum_{g=1}^G P(c_i = g | X_{ci}, \theta_G) \times \phi_{ig}(Y_i | c_i = g; X_{1i}, X_{2i}, Z_i, \theta_G) \end{aligned} \quad (1.45)$$

where p_{ig} is the probability of belonging to a latent class g and ϕ_{ig} is the density function of a normal distribution (marginal distribution of Y_i conditional on the random effects and latent classes) with mean $X_{1i}\beta + X_{2i}\nu_g + Z_i\mu_g$ and variance $V_i = Z_i B_g Z_i^T + \sigma_\epsilon I_{n_i}$.

The parameters can be estimated by maximizing the log-likelihood $l(\theta_G)$ defined as:

$$l(\theta_G) = \sum_{i=1}^N \log(L_i(\theta_G)) \quad (1.46)$$

Parameter optimization may be performed with the Marquardt algorithm of the Newton-Raphson method (See Section 1.3.1).

Choice of the number of latent classes

The number of latent classes cannot be estimated and has to be known before fitting the model through evaluation criteria. In the literature, the choice of the number of latent classes is calculated by the help of the Akaike's information criterion (AIC), the Bayesian Information Criterion (BIC), or the consistent AIC (CAIC) [20] [137] [148]. However, the

AIC tends to overestimate the number of classes whereas the BIC tends to underestimate this number in small sample sizes [137].

The number of latent classes that minimizes the BIC value is preferred [89] [15]. The BIC criterion is defined as

$$BIC(G) = -2L(\theta_G) + n_\theta \log(N) \quad (1.47)$$

where G is the number of latent classes, n_θ the number of estimated parameters and N the number of individuals.

Posterior probability and Classification

For a given individual, the posterior probabilities of belonging to each latent class may be calculated. The sum of these probabilities is equal to 1 ($\sum_{g=1}^G p_g = 1$).

The posterior probability for individual i of belonging to class g is defined by Bayes theorem as:

$$\hat{p}_{ig} = p_{ig}(c_i = g | X_i, Y_i, \hat{\theta}) = \frac{P(c_i = g | X_{ci}, \hat{\theta}) \times \phi_{ig}(Y_i | c_i = g, X_{1i}, X_{2i}, Z_i, \hat{\theta})}{\sum_{k=1}^G P(c_i = k | X_{ci}, \hat{\theta}) \times \phi_{ik}(Y_i | c_i = k, X_{1i}, X_{2i}, Z_i, \hat{\theta})} \quad (1.48)$$

where $\hat{\theta}$ is the set of estimated parameters with equation 1.45.

Then, each individual i is assigned to the latent class g that gives him the highest posterior probability of belonging \hat{p}_{ig} . The mean of the posterior probabilities in each class may help evaluating the quality of the discrimination ability of the fitted model. For example, if all individuals are classified in a class with a posterior probability equal to 1, the mean posterior probability will be equal to 1. This means that the classification performed with the model is perfect. The ability of the discrimination performed with this model weakens as the mean of the posterior probabilities approaches zero.

Prediction

Empirical Bayes estimates of random effects:

The Empirical Bayes estimates of random effects in the latent class mixed model is transformed from the linear mixed model defined in equation 1.40. So, the class-specific empirical Bayes estimate for the latent class models is defined as:

$$\hat{u}_{ig} = w_g^2 \hat{B} Z_i^T V_{ig}^{-1} (Y_i - X_{1i} \hat{\beta} - X_{2i} \hat{\gamma}_g) \quad (1.49)$$

where w_g is a proportional coefficient that allows for a class-specific intensity of individual variability ($w_g = 1$ for identifiability).

Marginal and subject-specific predictions

In the latent class mixed model, the marginal and subject-specific predictions can be calculated by replacing the fixed and random effect parameters by their estimated ($\hat{\beta}$, $\hat{\gamma}_g$, and \hat{u}_{ig}) values in equation 1.43.

When $G > 1$, the class-specific marginal and subject-specific predictions in the latent class linear mixed model are calculated as

$$\begin{aligned} \hat{Y}_{ijg}^{(M)} &= X_{1ij}^T \hat{\beta} + X_{2ij}^T \hat{\gamma}_g \\ \hat{Y}_{ijg}^{(SS)} &= X_{1ij}^T \hat{\beta} + X_{2ij}^T \hat{\gamma}_g + Z_{ij}^T \hat{u}_{ig} \end{aligned} \quad (1.50)$$

The class-specific subject-specific residuals are calculated with the average of the subject-specific predictions over all classes as:

$$\begin{aligned}
R_{ij}^{(M)} &= Y_{ij} - \sum_{g=1}^G \hat{p}_{ig} \times \hat{Y}_{ijg}^{(M)} \\
R_{ij}^{(SS)} &= Y_{ij} - \sum_{g=1}^G \hat{p}_{ig} \times \hat{Y}_{ijg}^{(SS)}
\end{aligned}
\tag{1.51}$$

Finally, the evolutions of the class-specific mean are calculated for each class weighting the mean of the subject-specific predictions by the class-membership probabilities as defined in equation 1.48.

Evaluation criteria

The performance of a latent class linear mixed model can be assessed with several evaluation criteria [37] such as:

1. BIC values
2. Comparison of the subject-specific predictions with the observed values
3. The mean of the posterior probabilities calculated at each class (as indicated above)
4. The proportion of individuals with a posterior probability located over a given threshold (e.g., 0.8 or 0.5 depending on the classification ambiguity).
5. Entropy measure as $1 - \frac{-\sum_{i=1}^N \sum_{g=1}^G \hat{p}_{ig} \times \log(\hat{p}_{ig})}{N \times \ln(G)}$. The discrimination ability is better when the entropy is close to 1.

1.4 Imaging

Magnetic Resonance Imaging (MRI) is a non-invasive medical imaging method that has a wide range of applications in medical fields to discriminate the healthy tissue from pathological tissue, in the diagnosis and follow-up of many diseases. Moreover, MRI has progressively gained in importance in neurosciences because of its ability to produce high quality images of brain and spinal cord tissue in any plane.

MRI can be divided mainly in two groups such as conventional and advanced imaging. Conventional imaging is the standard MRI techniques and has played an important role to observe anatomical changes in brain. Advanced imaging provides complementary physiological informations such as the detection of structural, functional and metabolic changes which are useful for early diagnosis and treatment monitoring.

MRI uses the body's natural magnetic properties to produce the images from any part of the body. The human body is mostly water. Water molecules (H_2O) contain hydrogen nuclei called proton. The protons are randomly aligned in three dimensions (on X, Y and Z planes) under normal conditions. When the body is placed in a strong magnetic field (B_0), such as an MRI, the protons' axes all line up like a small bar magnet on the plane parallel to B_0 (Z plane) (See Figure 1.7 and 1.8).

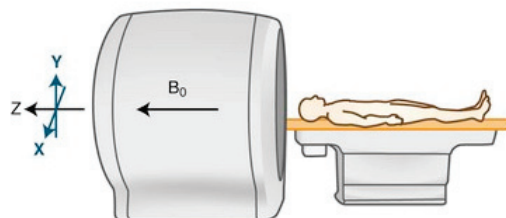


Figure 1.7: Magnetization in a MRI. Direction of external magnetic field is in the head-foot direction in the MRI.

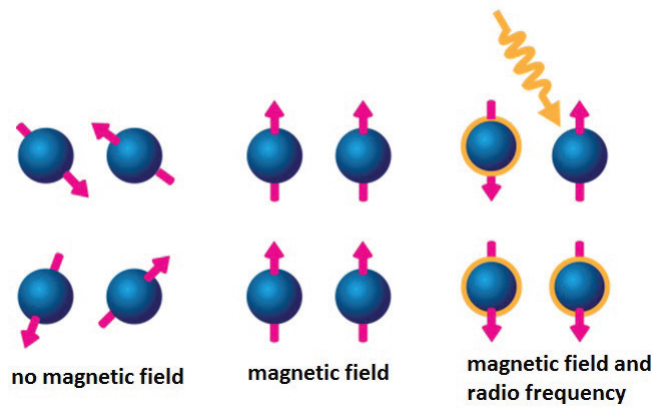


Figure 1.8: Effect of magnetic field and radio frequency on protons in the hydrogen molecule.

As a next step for obtaining the images, a radio frequency (RF) pulse is applied to the part of the body being imaged. This is an electromagnetic wave and RF causes the protons to change their orientation (See Figure 1.8) and to gain the energy. Once the source of RF is switched off, the protons return to their previous state quickly, i.e. relaxation. In relaxation process, the protons emit an amount of energy that is translated into images.

Fourier transformation is used to convert the energy frequency from each location to corresponding intensity levels, which are then displayed as shades of gray in the images. By varying the applied and collected RF pulses, different types of images can be created.

Relaxation time can be divided into two component such as the T1 and T2 relaxation time. T1 relaxation time is measured on the longitudinal (Z) plane and T2 relaxation time on the transverse (XY) plane. Relaxations time features are regularly used to show contrast between different tissues because each tissue has different T1 and T2 relaxation times.

One example of RF pulses is the 90-degree RF pulse that flips the entire net magnetization vector (summation of all the magnetic moments of the individual protons) by 90 degrees to the transverse plane (See Figure 1.9). When the pulse is turned off, two processes begin simultaneously: increase of the magnetization on longitudinal plane and decrease of the magnetization on transverse plane.

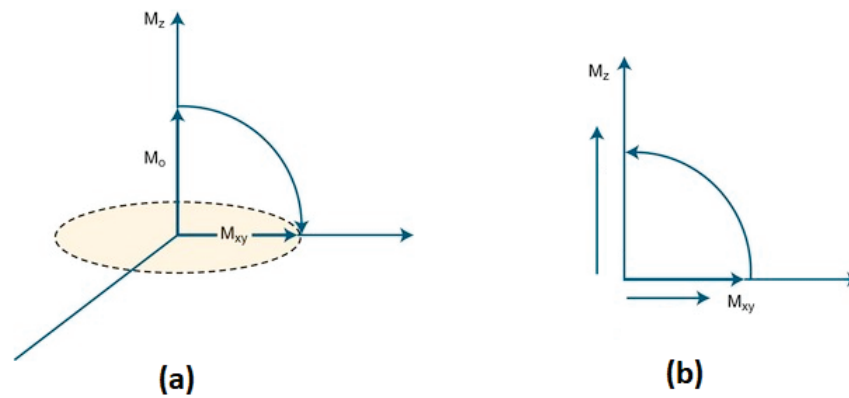


Figure 1.9: (a) When a 90-degree pulse is applied, the net magnetization vector of the protons is flipped from the longitudinal plane to the transverse plane. (b) When a 90-degree pulse is turned off, the net magnetization vector of the protons is flipped from the transverse plane to the longitudinal plane.

Most MRIs are in black/white with shades of gray : the white color refers high signal intensity, the grey color corresponds to the intermediate signal intensity, and the black color refers to low signal intensity. A **hyperintensity** is an area that appears brighter in color than the area we are comparing it to; a **hypointensity** would be darker in color.

1.4.1 Conventional Imaging

T1-Weighted Imaging

T1-Weighted Imaging (T1WI) uses differences in the T1 relaxation times of tissues and reflects the amount of time that the protons realign with the main magnetic field B_0 in longitudinal plane after RF pulses turn off.

T1WI shows a high spatial resolution and a high contrast between gray and white matter because of their different relaxation time. Chronic tissue damage is better seen in hypointense. Besides, T1-weighted images measure the volume and thickness of the cerebral cortex, which provides valuable information on neurodegenerative diseases as well as many other neurological diseases [99].

A contrast agent (called gadolinium) can be injected before the imaging process which improves the visibility of some pathologies. This imaging is called Contrast-Enhanced T1-Weighted Imaging and frequently used for Multiple Sclerosis patients to visualize the lesions clinically silent [13].

T2-Weighted Imaging

T2-Weighted Imaging (T2WI) highlights differences in the T2 relaxation time of tissues and reflects the amount of time that the protons realign with the main magnetic field in transverse plane. T2WI is used for characterizing abnormalities such as edema, tumors, inflammation, and white matter lesions [175] [79] [217].

Fluid-Attenuated Inversion Recovery

The fluid-attenuated inversion recovery (FLAIR) imaging corresponds to T2WI in hyperintension where the signal intensity of Cerebro Spinal Fluid (CSF) is suppressed. By using the FLAIR imaging, the periventricular hyperintense lesions such as multiple sclerosis lesions become more noticeable because CSF is suppressed [4]. For this reason, FLAIR is often preferred to T2WI because it better visualizes lesions in periventricular regions. In addition, FLAIR imaging is highly used in stroke to contour the boundaries of the irreversibly damaged brain tissue in stroke patients.

The difference between T1WI, T2WI and FLAIR is shown on Figure 1.10. The CSF appears dark on T1WI and FLAIR images, but brighter in color with T2WI. Moreover, T2WI and FLAIR show White Matter (WM) in dark gray, and Gray Matter (GM) in light gray. In addition, T1WI shows white matter in gray and white matter brighter in color.

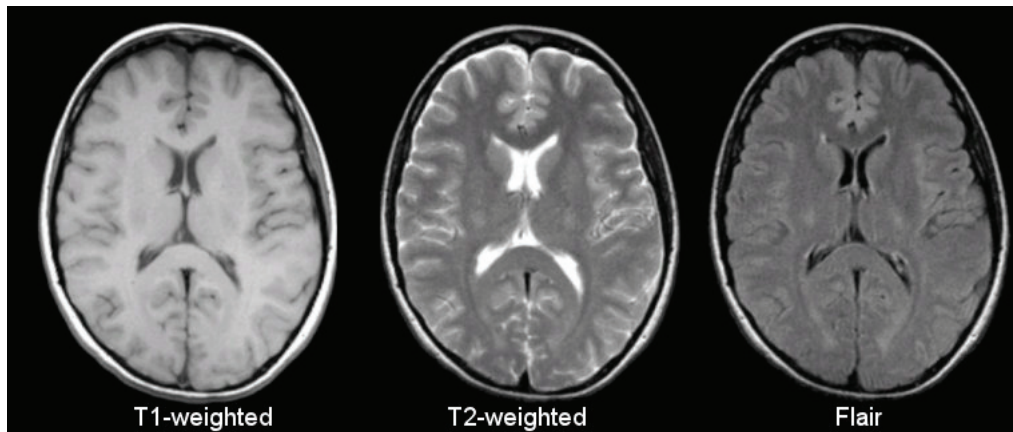


Figure 1.10: A brain imaging obtained respectively with T1-Weighted Imaging, T2-Weighted Imaging and FLAIR.

Perfusion-Weighted Imaging

Perfusion-Weighted Imaging (PWI) allows measuring hemodynamic circulation within the cerebral parenchyma after injection of a contrast agent (gadolinium). PWI is most commonly used with T2WI because of the linear relationship between the tissue contrast agent concentration and the change in the T2 relaxation time [176] [190]. The passage of the contrast agent causes a loss of signal because T2WI is exquisitely sensitive to intravoxel linewidth changes [228]. The tissue signal changes caused by susceptibility of T2WI create a hemodynamic time-concentration intensity curve presented in Figure 1.11.

The arrival of the contrast agent is calculated in each voxel of the scanned tissue using a voxel-specific curve fitting to an estimated gamma variate function (See Figure 1.11) [211]. PWI provides measurements of cerebral blood volume for a given amount brain tissue and the temporal delays of the first pass of gadolinium circulation [211] [241] such as:

- the volume of blood in a given amount of brain tissue: cerebral blood volume (CBV).
In addition, the area under the concentration-time curve gives the CBV.
- the volume of blood passing through a given amount of brain tissue per unit of time: cerebral blood flow (CBF)

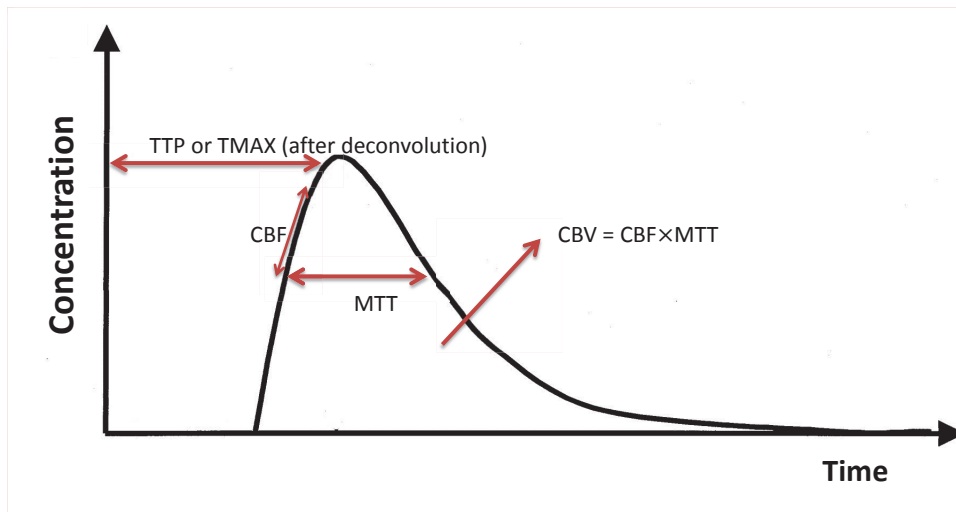


Figure 1.11: Contrast agent concentration change over time measured with T2WI in a brain voxel.

- the mean transit time (time taken to transit through the tissues by the contrast agent, $MTT = CBV / CBF$)
- the time to peak of the tissue concentration-time curve (TTP)
- the time to the peak of the residue function obtained by deconvolution (TMAX)

Diffusion-Weighted Imaging

Diffusion-Weighted Imaging (DWI) shows the random Brownian motion of water molecules within a voxel of cerebral tissue. Cerebral abnormalities leads to energy metabolism changes with the failure of the Na^+/K^+ and other ionic pumps [125]. This causes a transfer of water from the extracellular to the intracellular part causing a cytotoxic edema [7].

The reduction of water diffusion in the extracellular compartment is detected with DWI within minutes of vessel occlusion and the magnitude of diffusion (of water molecules) can

be assessed quantitatively with the apparent diffusion coefficient (ADC) [208] [171]. This assessment can be done using different acquisition parameters via changing gradient amplitude. The relationship between DWI, ADC parameters, and the acquisition parameter is given in equation 1.52:

$$DWI = S_0 \times \exp(-b \times ADC) \quad (1.52)$$

where b is the acquisition parameter and determines the strength and duration of the diffusion gradients and S_0 is signal intensity without diffusion weighting [168].

1.4.2 Advanced Imaging

Diffusion Tensor Imaging

Diffusion Tensor Imaging (DTI) is a very sensitive method in identifying the microstructural changes in the CNS that cannot be measured with conventional imaging [178]. DTI is based on a three-dimensional measurement of water mobility in three dimensions within tissues. Water mobility is equally distributed (isotropic). However, because of the highly organized myelinated axonal fiber tracts, the diffusivity of water is not the same in all directions (anisotropic) and water mobility is higher in one dimension than in the others (See Figure 1.12). Longitudinal diffusivity reflects at diffusion parallel to the fibers (axial diffusivity $\lambda_a = \lambda_1$), the mean of the two other diffusivities is called radial diffusivity ($\lambda_r = (\lambda_2 + \lambda_3)/2$) and the mean of all of three diffusivities is called Mean Diffusivity (MD) ($MD = (\lambda_1 + \lambda_2 + \lambda_3)/3$) and MD is independent from the orientation of the structures. Another dimensionless index of anisotropy of DTI is the fractional anisotropy (FA). FA takes values between 0 and 1: 0 for complete isotropy and 1 for completely anisotropy.

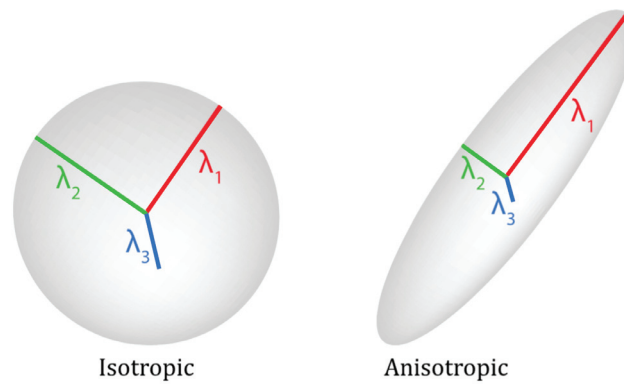


Figure 1.12: Figure on the left side shows the diffusion of the water in space (isotropic diffusion). On the right side, the figure shows the diffusion in brain (anisotropic diffusion). λ_1 : axial diffusivity, $(\lambda_2 + \lambda_3)/2$: radial diffusivity, $(\lambda_1 + \lambda_2 + \lambda_3)/3$: mean diffusivity.

Magnetic resonance spectroscopic imaging

Magnetic resonance spectroscopic imaging (MRSI) provides an important understanding of the in vivo chemical–pathological changes in neurological diseases measuring in normal appearing brain tissue [142] [179]. As MRSI can show the changes in different metabolites', characterizing different phases of a neurological disease is possible. The main metabolites are N-acetylaspartate (NAA), choline (Cho), creatine (Cr), lactate (Lac), glutamate and glutamine (Glx), and free lipids.

1.5 Overview of Stroke

1.5.1 Introduction

Stroke is considered as one of the most destructive neurological diseases and is the second most common cause of death and major cause of disability worldwide [53]. Stroke has also a negative impact on psychological and social situations such as problems with family relationships, deterioration in sex life, economic difficulties related to loss of the work, and deterioration in leisure activities [45].

Strokes are either ischemic or hemorrhagic, their managements are thus very different. Almost 85% of strokes are ischemic resulting from deterioration of blood flow by occlusion of an artery by an embolus from the heart or from acute thrombotic occlusion within the brain [10]. when the blood flow interruption is severe or lasts a long time, death of brain tissue occurs and an irreversibly infarcted zone is formed within the brain. This zone is called **ischemic core** (See Figure 1.13). When the blood flow is less severe, or the interruption occurs during only a short period of time, the brain tissue may recover partially or completely. This salvageable tissue is known as **ischemic penumbra** and is the target of the therapeutic interventions.

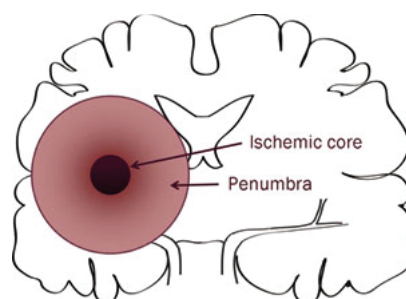


Figure 1.13: Ischemic core and penumbra after stroke. While neurons in the ischemic core are considered irreversibly damaged, neurons in the penumbra are salvageable and are potential targets for therapeutic interventions [43].

The National Institutes of Health Stroke Scale (NIHSS) is an assessment tool that provides a quantitative measure of stroke-related neurologic damage. The NIHSS is widely used to evaluate disability in patients with stroke, determine the appropriate treatment, and predict patient outcome. A practitioner evaluates the patient's ability to answer questions and perform activities considering the levels of consciousness, language, neglect, visual-field loss, extraocular movement, motor strength, ataxia, dysarthria, and sensory loss. The NIHSS ranges between 0 and 42: 0 for no stroke symptoms, 1-4 for minor stroke, 5-15 for moderate stroke, 16-20 for moderate to severe stroke, and 21-42 for severe stroke [207].

The risk factors for ischemic stroke can be classified into two subgroups: modifiable and unmodifiable. The modifiable risk factors are hypertension, diabetes, ischemic heart disease, atrial fibrillation, valvular heart disease, cigarette smoking. The unmodifiable risk factor is age. The risk of stroke doubles every decade above age 55 [107]. In developed countries, stroke affects usually people aged 70-75 years whereas stroke occurs in people less than 60 years in Sub-Saharan Africa [154] [41]. As hypertension causes a reduction in the external diameter of blood vessels; it has a powerful effect on the cerebral circulation [95]. For this reason, hypertension is the second strongest risk factor and people with hypertension are almost 3 or 4 times more likely to have a stroke [34].

1.5.2 Epidemiology

Stroke is the second leading cause of death worldwide with an annual mortality rate of 5.5 million people [128]. The rate of stroke changes geographically. The highest rate is observed in Eastern Europe, North Asia, Central Africa, and the South Pacific (See Figure 1.14) [94]. The high rate of stroke is partly related to poverty, thus lack of knowledge about stroke risk factors and its warning signs in low- and middle-income countries [59].

A recent study (in 2015) has indicated that the absolute number of people affected by stroke has substantially increased across all countries in the world over the same time

period between 1990 and 2013, even in developed countries [59].

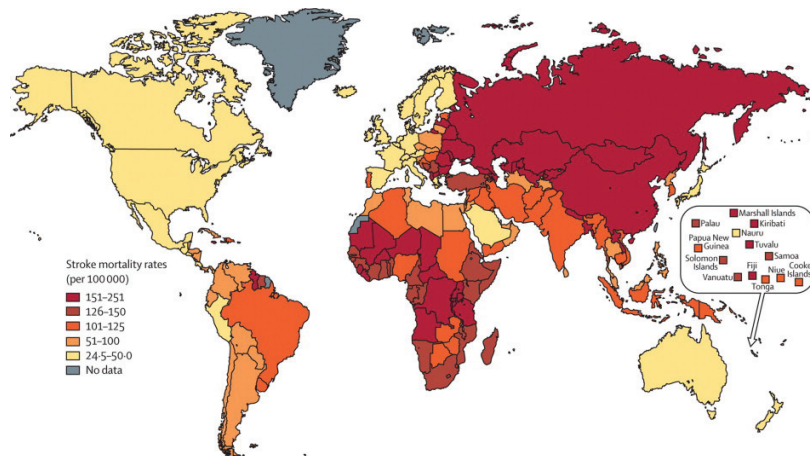


Figure 1.14: Global distribution of stroke mortality rates [107].

1.5.3 Pathophysiology

Cells of the Central Nervous System

A neuron is the basic unit of the nervous system. It mainly consists of

- a cell body that can be seen as the main part of the neuron where the nucleus is located,
- an axon surrounded by myelin sheaths and regular gaps (also called nodes of Ranvier) that helps transmitting electrical signals,
- dendrites extending from the cell body to receive the input elements to propagate to the cell body (See Figure 1.15).

The electrical signals are received by the dendrites from other neurons, pass through the cell body and then propagate by the axon to another neuron by synapses where the axon of one neuron meets the dendrite of another. This structure shows the unidirectional signaling of neuron mechanism [109].

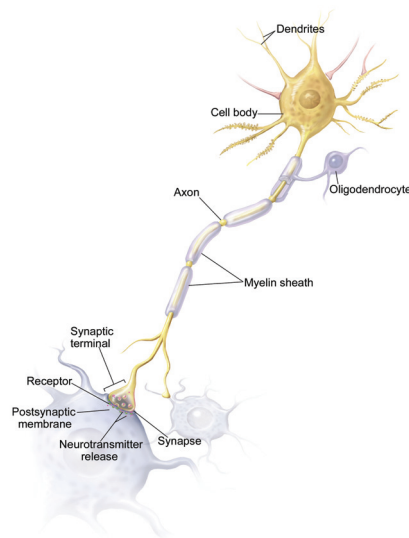


Figure 1.15: A neuron in the Central Nervous System.

The myelinated axons are found in the CNS as bundles and the structure of these bundles is called the "white matter" because of the white colored fatty substance (myelin sheaths). The other parts of the neuron such as cell bodies, dendrites, and unmyelinated axons form a structure called "gray matter" of the CNS. Due to saltatory conduction characteristic of a myelinated neuron, the white matter allows a quick transmission of signals between different areas of the gray matter within the CNS.

Pathophysiology in stroke

The future of brain cells located in the ischemic penumbra depends on changes in the blood flow. A brain cell needs about 50ml/100g/min blood flow with an oxygen metabolic rate of 3.5cc/100g/min [30] [52]. When the blood flow decreases below 10ml/100g/min, brain-cell functions are severely impaired. Moreover, neurons are unable to survive longer when the blood flow is below 5ml/100g/min [30] [52]. In complete absence of blood flow, the tissue becomes infarcted in 2-3 minutes.

If the blood flow is not restored, an ischemic cascade starts with an accumulation of sodium, calcium, and water inside the damaged neurons and other cells (in 3-24 hours).

This leads to the release of excitatory neurotransmitters that cause further deterioration of the blood-brain barrier (BBB) (in 2 days - 2 weeks) [52] [132]. The latter deterioration allows water to come out of the blood vessels. Thus, the net volume of water in the ischemic core increases causing vasogenic edema.

1.5.4 Imaging of Stroke

The first step after patient admission is distinguishing ischemic stroke from hemorrhagic stroke. This information is needed for an appropriate treatment of all patients with stroke.

The imaging tools help in diagnosis and patient follow-up. Computed tomography (CT) and MRI are the most frequently used imaging tools for patients with stroke. However, MRI has more advantages than CT though its acquisition time is longer. One of the advantages of MRI is its better ability to distinguish hemorrhage from thrombus than CT imaging [232]. Therefore, here, we will use MRI parameters.

DWI provides a marker of irreversibly damaged areas whereas PWI is recommended to characterize the status of brain tissue blood supply and to identify the reversible area in ischemia [205] [240] [91]. However, there is a high discrepancy between the lesions identified by DWI vs. PWI while the performances of DWI and PWI data combined has been shown superior to that of DWI data or PWI data alone in identifying the tissue at risk of infarction [205] [240].

Figure 1.16 shows the DWI hyperintensity and ADC hypointensity images in a patient with stroke. The areas with high signal are readily apparent and the infarcted area is more visible with DWI parameter than with ADC parameter.

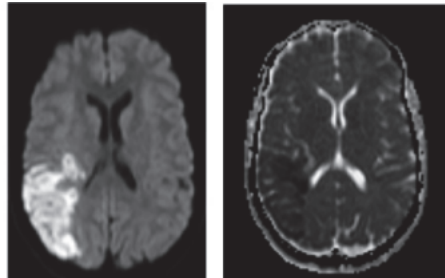


Figure 1.16: MRI of the brain. DWI (image at left side), and ADC map (image at right side), showing high signal intensity on the DWI and low signal intensity on the ADC parameter in a patient with ischemic stroke [139].

FLAIR imaging at one month after stroke onset is commonly used to locate the boundaries of the irreversibly damaged brain tissue in patients with stroke [21]. Figure 1.17 provides T2WI, FLAIR and DWI parameters in a patient with ischemic stroke. The T2WI and FLAIR show similar hyperintense signals. However, the infarcted core close to ventricular regions is better identified with FLAIR. DWI shows the infarcted area in hyposignal and reveals no evidence of associated ischemia.

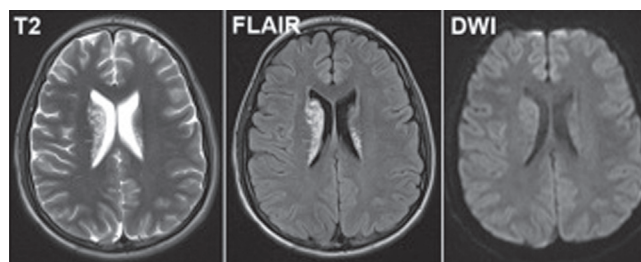


Figure 1.17: T2, FLAIR and DWI imaging in ischemic stroke [81].

1.5.5 Disease-modifying therapies

The ischemic penumbra can be observed as a mismatch between a large PWI lesion and a smaller DWI lesion (See Figure 1.18). Figure 1.19 shows also the mismatch observed with DWI and PWI parameters on the human brain. The images show that the infarcted core observed on DWI is smaller than the reversibly damaged area shown with PWI. The ischemic penumbra can be observed in at least 80% of patients within 3 h of stroke onset, but the number of patients with penumbra decrease with time [177]. The probability of penumbra' infarction is related to the severity and duration of the hypoperfusion [27].

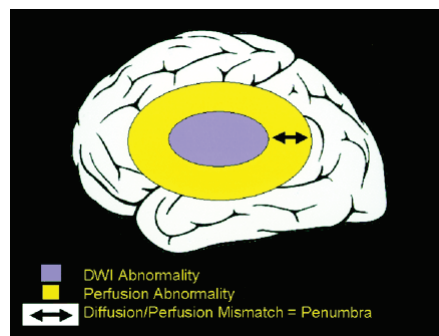


Figure 1.18: DWI/PWI mismatch [114].

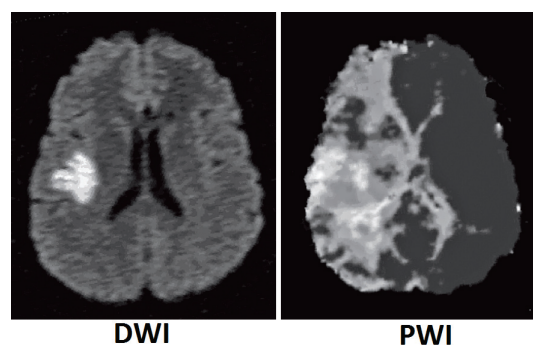


Figure 1.19: MRI of the ischemic penumbra in a patient with stroke [53].

The ischemic penumbra is the target of therapy and DWI/PWI mismatch is highly used to identify patients who are the most likely to benefit from thrombolytic treatment beyond 3 hours after an ischemic stroke [53]. Treatments to re-open the blocked blood vessel and

reperfuse the brain are available. Intravenous recombinant tissue plasminogen activator (tPA) is the only licensed therapy for treatment of acute ischemic stroke and the use of this treatment within the first 3 hours after stroke onset is recommended [231] [218]. Its effectiveness is improved when applied but the benefits decrease rapidly over the first few hours after stroke onset.

The aims of tPA are arterial recanalization and recovering of the ischemic penumbra, a viable but critical region around the irreversibly damaged infarct core [214]. The use of tPA is associated with tissue reperfusion and a decrease in infarct growth, presence, and magnitude of the ischemic penumbra after a certain period of time [158] [121].

Meta-analyses have indicated that tPA may be effective in a smaller portion of patients beyond 3 hours [174] [83]. Approximately 2% to 5% of patients with acute ischemic stroke receive tPA because this drug may cause intracranial hemorrhage and other troubles. For example, one of the meta-analyses has shown that symptomatic intracranial hemorrhage, major systemic hemorrhage, and angioedema have been observed in approximately 6%, 2%, and 5% of patients, respectively, of patients treated with tPA.

1.6 Overview of Multiple Sclerosis

1.6.1 Introduction

MS is a chronic inflammatory disease characterized by demyelinating plaques and neurodegeneration within the central nervous system (CNS) [32] [103]. The immune system attacks the myelin sheath that covers the nerves and then gives rise to clinical abnormalities. The etiology of MS is still unknown, but it has been believed that the disease is triggered by environmental factors in a person who is genetically predisposed [209] [167].

The active inflammatory phase in the CNS can cause relapses defined as worsening of existing symptoms or appearance of new symptoms that last more than 24 hours and happen at least 30 days after a previous relapse. Clinical worsening can be observed between the relapses but many patients recover from their relapses without worsening.

The length, severity, and number of relapses, as well as clinical disability evolution are highly heterogeneous between patients. Therefore, a powerful long-term prediction of relapses or clinical evolution is currently very difficult.

Moreover, there is no stopping disease therapy for multiple sclerosis. Disease-modifying drugs, including immunotherapies, do exist and help reducing autoimmune disorders, thus, speed recovery from relapses, reduce inflammation. These drugs may then change the course of the disease and manage symptoms.

1.6.2 Epidemiology

MS is the most common non-traumatic disabling neurological disease of adulthood. It occurs generally between ages 18 and 40 years [38]. The estimated global prevalence of MS is approximately 30/100 000 and there were at least 1.32 million patients with MS according to the 2008 MS Atlas [26]. Even the etiology of MS is still largely unknown but there are evidence for roles of both environmental and genetic factors (ultraviolet rays, vitamin D levels, cigarette smoking, and migrations) [209] [167].

Environmental factors

There are large differences in the reported prevalence of MS, both within and between countries (See Figure 1.20). Disease frequency changes depending on the geographical areas, especially the latitude, consequently to low exposure to ultraviolet radiation (UVR) and low vitamin D levels. Intensive sun exposure during childhood and early adolescence are associated with a reduced risk of multiple sclerosis whereas the people with low vitamin D level have higher risk of developing MS [225] [98] [108] [147] [131]. Moreover, low vitamin D levels increase the relapse risk [203] [202] [201]. Northern Europe, Northern United States, Southern Canada, New Zealand, and Southern Australia have the highest risk (prevalence: 80/100 000). Southern Europe, South America, and the rest of Australia have a medium risk (prevalence: 30-80/100 000). The lowest risks (prevalence: 30/100 000) are in Africa and Asia [104] [135].

France is in an area with a medium to high prevalence of MS. French farmers account for only 7% of the French population and the prevalence of MS changes significantly between the northeastern and the southwestern regions [230] (See Figure 1.21). Another study found no north-south gradient for MS prevalence but confirmed the high prevalence in the North-Eastern France in a large population representative of the French population [71].

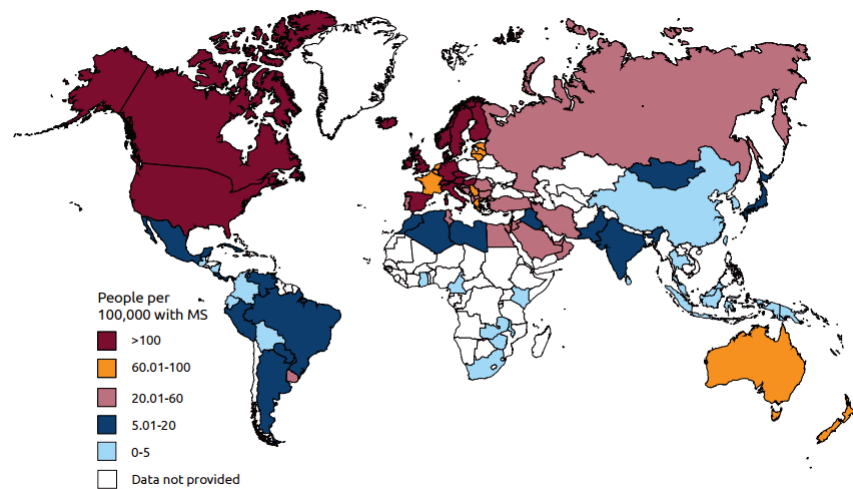


Figure 1.20: The estimated number of people with MS in 2013 in each country.

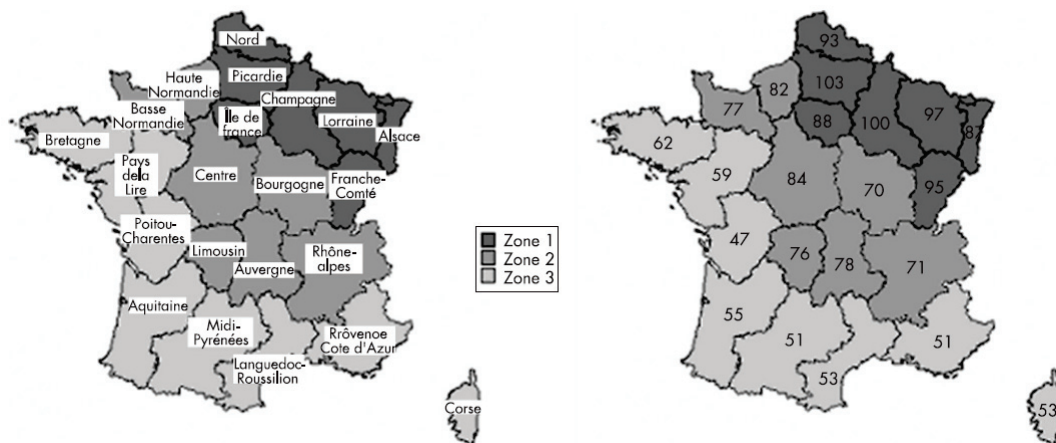


Figure 1.21: Estimates of the regional prevalence of multiple sclerosis on 1 January 2003, per 100 000 inhabitants, standardised by age, among French farmers [230].

Genetic factors

The heterogeneous prevalence of MS in different regions cannot be explained by one or a few known environmental factor(s). Major differences in prevalence are observed between countries located at the same latitude. Moreover, similar prevalences occur in regions with individuals of same ethnic origin [184]. It is suspected that there is a role for susceptibility genes in the development of MS [146] [237].

The familial occurrence of MS has an effect because MS prevalence increases with the increase of kinship degree. The big concordance rate in monozygotic twins (20-30%)

versus dizygotic twins (3-5%) indicates a major genetic etiologic contribution [212] [55] [184].

Many genetic studies have shown the genes that encode for the major histocompatibility complex (MHC) (also known as the human leukocyte antigen (HLA) system) have the strongest effect on MS among genetic factors [1] [152] [11] [243].

Women are also twice more affected than men in most of studies. In Europe, the ratio ranges from 1.1 to 3.4 [2] [55] [118].

1.6.3 Physiopathology

MS is an inflammatory, demyelinating, and neurodegenerative disease of the CNS [66]. The disease presents first as an inflammation of the CNS provoked by T-cells, B-cells, and macrophages. Inflammatory can provoke then the neurodegeneration; demyelination and axonal loss form then lesions (also called plaques) in the white and/or the grey matter [151] [92] [183]. The lesions are located within the brain, primarily in myelin-dense white matter regions including the corpus callosum, the spinal cord, and then optic nerves. Further, lesions may lead to brain atrophy [66] [70] [57]. Many white matter lesions are detected around the ventricles but lesions may also occur in the deep white matter far from the ventricles as well as in the grey matter [123].

Lesions in Corpus Callosum

As indicated above, MS affects especially the white matter of the brain and the lesions appear mostly around the ventricles and within the periventricular white matter. The floor of the lateral ventricles is formed by a region named Corpus Callosum (CC) (See figure 1.22), one of the most frequently affected regions: in 93% of Patients with MS, lesions were found in CC [14] [77] [74]. Moreover, CC is the major myelinated bundle of the brain and provides the connection between cortical and subcortical regions of brain

hemispheres. For this reason, the effects of lesions in CC are more severe versus lesions in the lobar white matter.

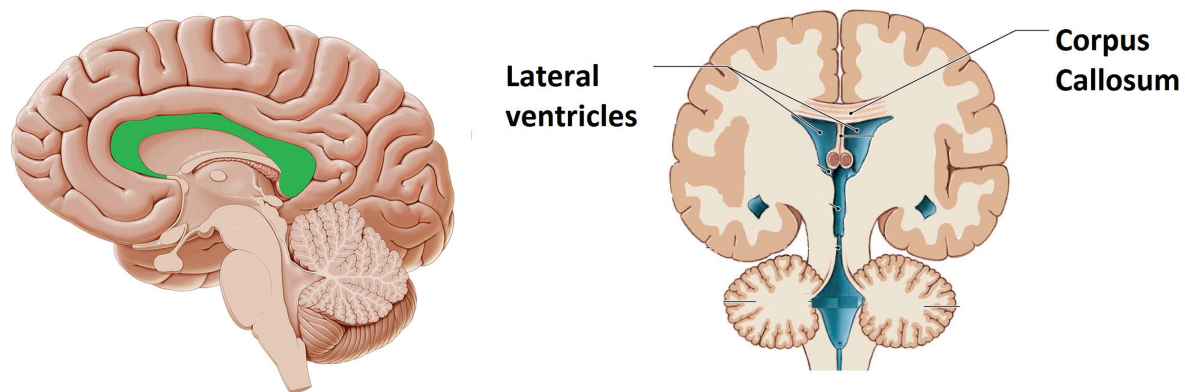


Figure 1.22: Figure at the left side: A brain image at sagittal plane where the green part shows the Corpus Callosum. Figure at the right side: A brain image at coronal plane indicating the lateral ventricles and corpus callosum.

Several studies demonstrated that callosal changes measured with Diffusion Tensor Imaging (DTI) were correlated with cognitive and physical disability [242] [173] [198] [127]. The association between clinical scores and CC atrophy measured with conventional imaging is reported in some studies [192] [242] [156].

As a conclusion, the biophysical and biochemical changes in all the CNS -due to the lesions- may give rise to clinical symptoms or signs. As the density of these alterations and their location in the CNS differs from one patient to another, the disease evolution is also highly heterogeneous among patients with MS.

1.6.4 Evolution and severity of clinical disability

Evolution of clinical disability

The length, severity, number of relapses, and clinical disability evolution are highly heterogeneous between patients. Thus, patients with MS do not present a homogeneous population and they are classified into four clinical subtypes according to the disease evolution: Clinically Isolated Syndrome (CIS), Relapsing-Remitting (RR), Primary Progressive (PP), and Secondary Progressive (SP) [129] (See Figure 1.23).

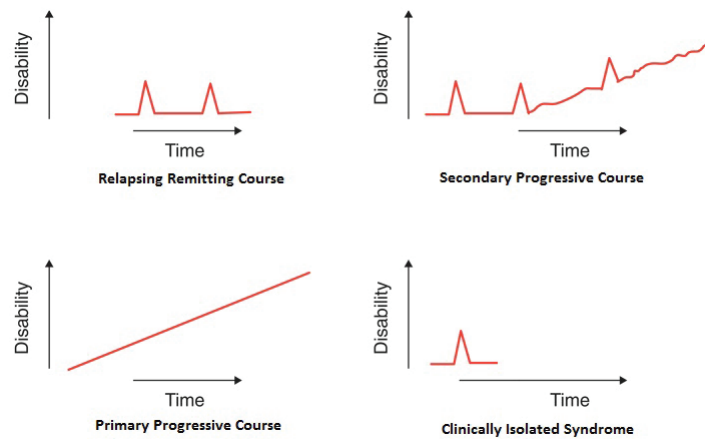


Figure 1.23: Change in disability score over time in the four MS clinical subtypes: Relapsing-Remitting (RR), Secondary Progressive (SP), Primary Progressive (PP), Clinically Isolated Syndrome (CIS).

Clinically isolated syndrome (CIS) patients show the first clinical episode such as optic neuritis, brainstem syndrome, or partial myelitis. This syndrome is known as the first phase of RR and should be followed-up to determine the subsequent disease evolution. Several studies have shown that the majority of patients shift toward RR within two or three years of follow-up [110] [143].

A relapse (also called attack) is defined as the appearance of new clinical symptoms or signs lasting at least 24 hours with a one-month interval before the next attack. Most patients (85%) initially have RR characterized by discrete relapses with full or partial recovery accompanied by little or no disability evolution. However, over time RR can

progress to SP where the frequency of relapses decreases and the disability progression is observed between relapses [129]. Approximately 65% patients with RR develop SP within 5-15 years of symptom onset [182] [187]. A minority of patients (15%) have PP where the irreversible progression has occurred from the beginning in the absence of relapses [220]. PP is considered to be a mostly noninflammatory subtype [130] [38].

The inflammation phase in the white matter is mostly observed among RR patients and the focal inflammatory event in the CNS is the main cause of relapses [229]. The inflammation may lead to axonal loss then to brain atrophy which is more frequently observed among patients with progressive evolution (PP and SP) [123]. Brain atrophy can be seen as the result of the transition from RR to SP patients and associated with irreversible disability in progressive evolution (See Figure 1.24) [223] [19].

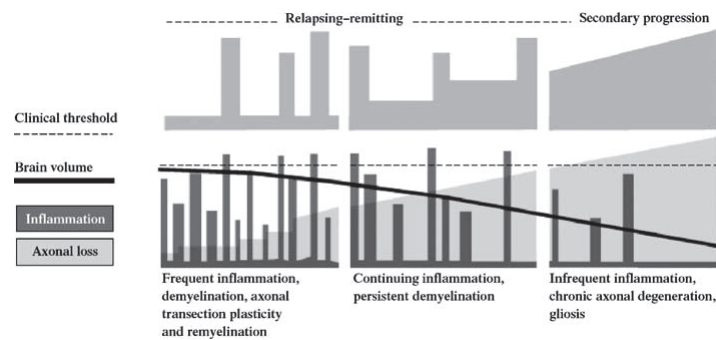


Figure 1.24: Inflammation and axonal loss evolution during the transition from RR to SP. Relapse frequency decreases giving rise to axonal loss followed by permanent disability [229].

Severity of clinical disability

The Expanded Disability Status Scale (EDSS) is the most frequently used clinical severity score to quantify disability in MS. The EDSS scores range from 0 to 10 by 0.5 unit increments where 0 indicates no disability and 10 death due to MS (See Table 1.2). EDSS scoring between 0 and 4.5 uses impairment measures in eight functional systems (FS):

1. pyramidal (motor function) - weakness or impairment moving limbs

2. cerebellar - ataxia, loss of coordination or tremor
3. brainstem - nystagmus and impairment of speech and swallowing
4. sensory - loss of sensations
5. bowel and bladder function
6. visual function
7. mental functions
8. other

The functional systems are scored between 0 and 5 or 6 : 0 indicates no disability and 5 or 6 indicate most severe disability.

EDSS scores 0.5 to 2.0 refer to patients with minimal-clinical disability, 2.5 to 4.5 refer patients who are able to walk with increased limitation in walking but insufficient to constrain from normal activities of daily living or work with some impairments of FS , 4.5 to 6.0 refer to patients who need assistance for ambulation. Finally, EDSS scores 7.0 refer to patients restricted to wheelchair and EDSS score 9.0 to helpless-bed patients.

The EDSS score has been criticized for some limitations: (1) it focuses on walking as the main measure of disability because it is a mixture of impairment and disability, (2) its non-linearity because the progression between 1 and 5 is faster than between 5 and 7 [238] [196].

MS progression may also be scored with the multidimensional score, the Multiple Sclerosis Functional Composite (MSFC) [180]. The MSFC combines three timed tests into a single score and includes a timed 25-foot walk (T25FW), the nine-hole peg test (9HPT), and the Paced Serial Addition Test (PASAT). T25FW is averaged over 2 consecutive walks and reported in seconds. The 9HPT measures the manual dexterity and ask participants to repeatedly place and remove nine pegs into nine holes, one at a time, as quickly as

possible. The 9HPT is passed with both the dominant and non-dominant hand twice and is scored in seconds. The PASAT is a challenging task that involves working memory, attention, and arithmetic capabilities. A number is given every 3 seconds and the patient is asked to add it to the number he/she has just heard before. PASAT is scored between 0 and 60; the total number of correct answers is 60.

Score	Description
1.0	No disability, minimal signs in one FS
1.5	No disability, minimal signs in more than one FS
2.0	Minimal disability in one FS
2.5	Mild disability in one FS or minimal disability in two FS
3.0	Moderate disability in one FS, or mild disability in three or four FS. No impairment to walking
3.5	Moderate disability in one FS and more than minimal disability in several others. No impairment to walking
4.0	Significant disability but self-sufficient and up and about some 12 hours a day. Able to walk without aid or rest for 500m
4.5	Significant disability but up and about much of the day, able to work a full day, may otherwise have some limitation of full activity or require minimal assistance. Able to walk without aid or rest for 300m
5.0	Disability severe enough to impair full daily activities and ability to work a full day without special provisions. Able to walk without aid or rest for 200m
5.5	Disability severe enough to preclude full daily activities. Able to walk without aid or rest for 100m
6.0	Requires a walking aid - cane, crutch, etc - to walk about 100m with or without resting
6.5	Requires two walking aids - pair of canes, crutches, etc - to walk about 20m without resting
7.0	Unable to walk beyond approximately 5m even with aid. Essentially restricted to wheelchair; though wheels self in standard wheelchair and transfers alone. Up and about in wheelchair some 12 hours a day
7.5	Unable to take more than a few steps. Restricted to a wheelchair and may need aid in transferring. Can wheel self but cannot carry on in standard wheelchair for a full day and may require a motorized wheelchair
8.0	Essentially restricted to bed or chair or pushed in a wheelchair. Maybe out of bed itself much of the day. Retains many self-care functions. Generally has effective use of arms
8.5	Essentially restricted to bed much of day. Has some effective use of arms retains some self-care functions
9.0	Confined to bed. Can still communicate and eat
9.5	Confined to bed and totally dependent. Unable to communicate effectively or eat/swallow
10.0	Death due to MS

Table 1.2: The Expanded Disability Status Scale (EDSS) [105].

1.6.5 Imaging in MS

MRI is an important tool and the most extensively used technique in diagnosing and monitoring MS. MRI has a high sensitivity and a high specificity in detecting white and gray matter abnormalities even in case of clinically silent lesions of the CNS. The abnormalities observed on MRI reflect the underlying pathology: inflammation and axonal degeneration [112] [22]. Conventional imaging parameters are essential and acquired in all patients with MS: they have a high sensitivity in detecting lesions. In addition, the advanced imaging parameters are highly specific in detecting structural, functional, and metabolic changes which can not be observed with conventional imaging [62] [56].

Conventional imaging in MS

Conventional MRI has been found very promising in MS because:

1. they assist MS diagnosis in showing lesion dissemination over time and space
2. they exclude alternative disease probabilities (for example: Lyme disease, Lupus, Sjögren's syndrome) even at the earliest stages
3. they are high sensitive in showing CNS damage in MS in comparison with other neuroimaging techniques
4. they are able to provide objective metrics to help observe disease evolution and assess treatment efficiency
5. they help to identifying the lesions and their location

The most frequently used conventional MRI techniques to assess MS evolution are T1WI, contrast-enhanced T1WI, T2WI and FLAIR [65] [200]. T1WI and T2WI have important specificities and sensitivities in detecting tissue damage and to provide valuable predictive information on the evolution of MS [62] [67].

T1-Weighted Imaging The hypointense lesions (also called black holes) shown with T1WI were significantly correlated with disease progression and disability [224].

The early phase of neurodegeneration, the inflammatory edema, can be better observed in hyperintense with contrast-enhanced T1WI. Gadolinium is injected before imaging acquisition and it improves the visibility of inflammation as ringlike structure in T1WI (See Figure 1.25). Normally, gadolinium cannot pass from the bloodstream into the brain or the spinal cord because of a protection layer: the blood-brain barrier (BBB). However, the lesions may disrupt the BBB allowing gadolinium to pass through. Contrast-enhanced T1WI is thus able to detect clinically silent lesions, 5 to 10 times more frequently than clinical evaluation of relapses [13].

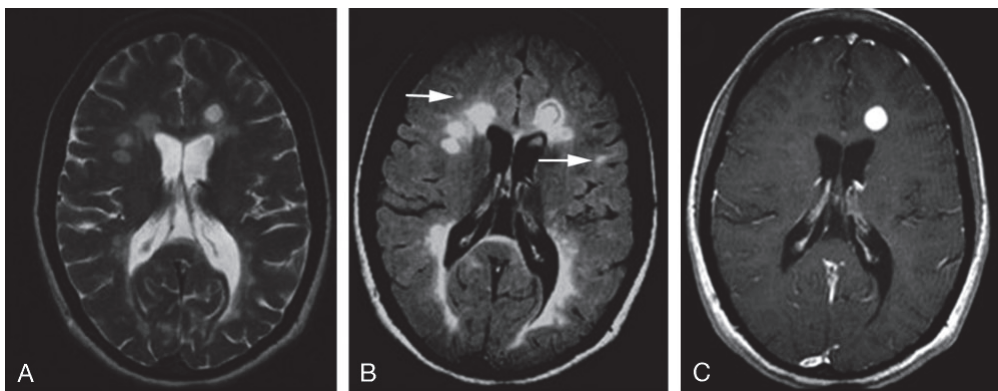


Figure 1.25: T2WI (A), FLAIR (B), and contrast-enhanced T1WI in a 30-year-old woman RRMS [74].

T2-Weighted Imaging In MS, T2WI has a great sensitivity in detecting lesions. The hypointense black-holes on T1WI appear as hyperintense MS lesions on T2WI.

T2-weighted hyperintense lesions' location has a great importance helping diagnosis and in differentiating MS from other diseases. MS lesions are mostly observed in periventricular, juxtacortical, and infratentorial regions [185]. Mc Donald criteria (2010 corrected version) underline the importance of the use MRI, especially T2-weighted lesions. Mc Donald criteria require a minimum of one T2 lesion in at least 2 of the following areas : periventricular, juxtacortical, infratentorial, and spinal cord (See Table 1.4) [161] .

Some cross-sectional studies have shown that brain atrophy observed with T2 imaging is significantly associated with physical disability [9] [8] [206]. Moreover, the baseline gray matter T2 hypointensity seems to be a strong predictor of the rate of whole brain atrophy over 2 years in RR patients who used the placebo because interferon may cause the brain atrophy [17].

FLAIR The suppression of the CSF helps to detect the lesions close to the CNS, for example, the periventricular part of the CNS. Some lesions can be detected by T1 or T2WI but can appear with FLAIR. FLAIR imaging is also the best way to identify lesions located in corpus callosum that cannot be seen by T2WI (See Figure 1.26) [4], thus, may help to differentiate MS from cerebrovascular diseases [74] .

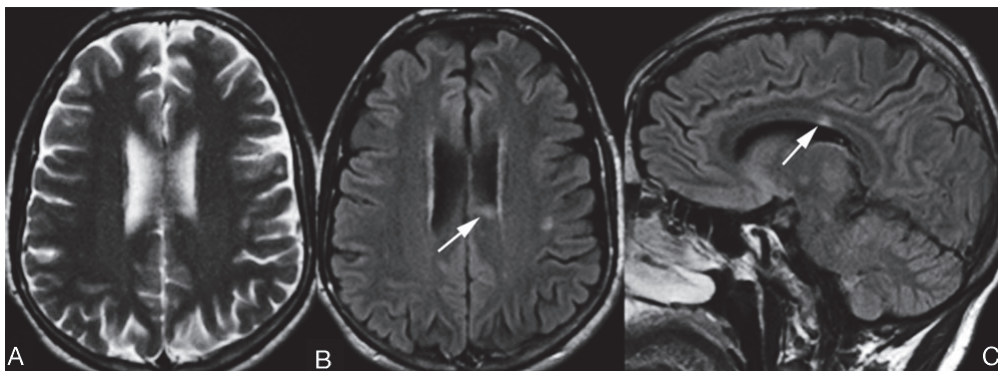


Figure 1.26: MS lesion in corpus callosum on T2WI (A), FLAIR (B) and FLAIR in sagittal plane (C) [74].

However, in FLAIR imaging, the contrast between the gray and the white matter is low, thus FLAIR has a relatively low sensitivity in detecting lesions located in the brainstem or the cerebellum, and performs less well in the posterior regions [172].

Advanced imaging in MS

In MS research, advanced imaging has demonstrated higher degrees of sensitivity and specificity in detecting inflammation and neurodegeneration than conventional imaging [62] [56] because they provide complementary information such as diffusion alteration and metabolic information in MS [61] [62] [63]. DTI and MRSI are the two most frequently used advanced techniques in MS. DTI is able to separate MS subtypes because the patients in different subtypes display different diffusivity patterns [56]. MRSI imaging enables quantifying metabolic abnormalities and leads to a better understanding of the processes that occur in lesions.

Diffusion Tensor Imaging DTI parameters inform about myelin and axonal loss as well as about the degree of inflammation. The measurements of λ_r and λ_a are significantly correlated with, respectively, myelin and axonal content [204]. MD and FA are primarily influenced by the free space caused by axonal and myelin loss and there is a significant difference between myelinated and non-myelinated nerves [93] [236] [60]. The presence of demyelination causes more space in the brain that increases MD and decreases FA measurements. Moreover, MD has been found useful for distinguishing different levels of MS lesions, which may be linked to various grades of clinical disability [189]. It has been shown that there is a significant difference in DTI parameters such as λ_a , λ_r , MD and FA between PP patients and control group [164] [199].

Several studies have demonstrated that callosal changes measured with DTI were correlated with cognitive and physical disability [242] [173] [127]. FA measurements were significantly lower in rostrum, body and splenium part of CC in patients with MS versus control group where MD was significantly greater [87] [181]. The association between EDSS scores and CC atrophy as measured by conventional MR is indicated in some studies [192] [87] [233] [242]. However, no significant correlation was found between disability and callosal atrophy as measured by conventional MR in RR patients [12]. These contrasts can result from insufficiency of conventional imaging of CC; the use of advanced

MRI techniques such as DTI can be more suitable for CC in patients with MS (See Figure 1.27).

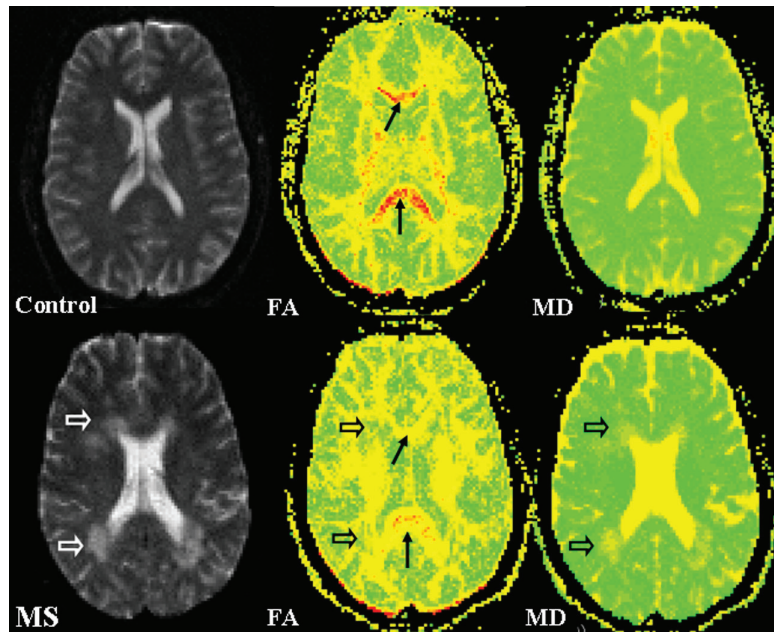


Figure 1.27: Images from a control subject (top row) and a patient (bottom row): DTI raw image (left column), color-coded FA map (middle column), and MD map (right column). FA takes the highest value (marked with red color) in the control subject whereas there is less red color in patients with MS, this indicates neurodegeneration. The lesions around CC are more visible on FA and MD maps [75].

Magnetic resonance spectroscopic imaging NAA is known as a neuronal marker of persistent neuronal/axonal damage [6] [210] [18]. The decrease in NAA indicates neuronal/axonal dysfunction or loss [5] [186]. Cho, Cr, Glu, Lac, and lipids inform about the pathophysiology of the inflammatory phase and tissue repair mechanisms in MS. Moreover, an increase in Cho:Cr ratio was found significantly related to acute lesion and inflammation which are observed in gadolinium injected T1-weighted lesions (See Figure 1.28) [239] [122].

In active inflammatory MS lesions, an increase in Cr and Cho and a decrease in NAA value have been observed. Also, in chronic lesions, NAA concentration decrease associated with axonal damage [31]. Changes in NAA and Cho inform about axonal damage and were found more correlated with the EDSS score than conventional MRI at an early stage of the disease [49]. Another study showed that the axonal damage assessed by NAA:Cr ratio decrease has a stronger correlation with clinical disability than T2-weighted lesion volume [72].

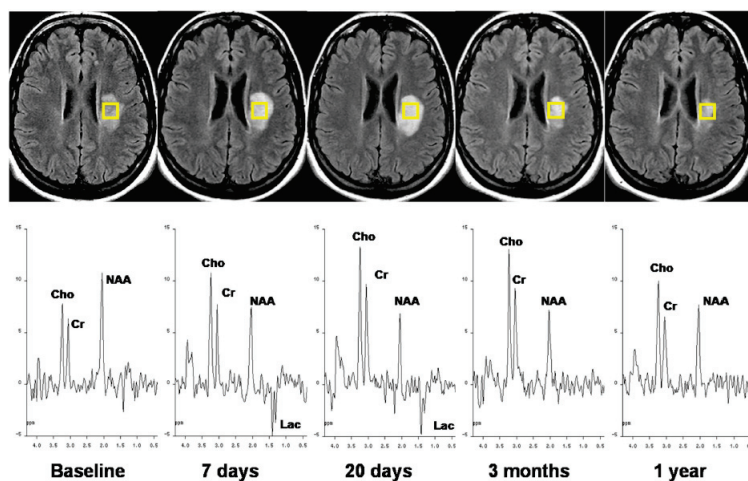


Figure 1.28: Serial T2-FLAIR images (top row) and spin-echo MR (bottom row) recorded from baseline to one year in acute multiple sclerosis lesion. FLAIR images show the evolution of lesion size that increases during the first 20 days and decreases afterwards. The spectroscopic graphs show a decrease in NAA from baseline up to 20 days and then stabilization up to 1 year. An increase in Cho during the first weeks followed by a partial recovery is also observed whereas Cr stays relatively stable at all time points [179].

1.6.6 Diagnosis

Clinical, imaging and laboratory exams help clinician confirm the diagnosis of MS. Patients with MS present a wide variety of clinical signs and symptoms such as fatigue, pain, depression, ataxia, bladder, bowel or sexual dysfunction, spasticity, sensory, visual, and cognitive impairments [169] [195] [145] [46] [153].

In addition to clinical examination, MRI is effective in diagnosing and monitoring the disease and also in assessing the effect of treatments on relapses in clinical trials [136] [222] [64]. Generally, one key element in MS diagnosis is the assessment of its "dissemination in space" (DIS, the lesions are observed in more than one region of the CNS) and "dissemination over time" (DIT, disease occurrence at different times: at least 30 days between two events). The sensitivity of MRI in identifying focal white matter abnormalities and clinically silent lesions plays an important role in evaluating DIS and DIT. Together with a contrast agent, MRI allows identifying inflammatory lesions as active and non-active [235] [222].

In the beginning, Schumacker criteria were used for diagnosis with DIS and DTI assessments of neurological impairments in different regions of the CNS (See table 1.3) [194]. Then, Poser criteria included CSF evaluations and paraclinical data as supplementary criteria to identify whether a person had possible, probable, or definite MS [163]. The criteria were revised in 2001, 2005, and 2010 and called "McDonald criteria". The latter incorporate clinical, paraclinical, and imaging data that help diagnosing MS in patients with a variety of disease presentations (See Table 1.4) [136] [162] [161]. These criteria underline the importance of MRI, especially T2-weighted lesions in spinal cord imaging, in the early diagnosis of MS [162] [161]. The reform of the latter criteria excluded the use of a contrast agent, saving thus time and cost.

Requirements for the diagnosis of clinically-definite MS
<ol style="list-style-type: none"> 1. Objective abnormalities on neurological examination attributable to dysfunction of the CNS. 2. Historical or objective evidence of the involvement of two or more separate parts of the CNS. 3. Objective signs of CNS disease should be attributable the white matter, with more than minor gray matter involvement disqualifying. 4. One of the following temporal patterns: <ol style="list-style-type: none"> a. Two or more episodes of worsening (relapse); each lasting at least 24 h, separated by one month or more. b. Slow or step-wise progression of signs and symptoms over at least 6 months. 5. Patients aged 10 – 50 years. 6. No better explanation for patient’s signs and symptoms.

Table 1.3: Schumaker criteria for the diagnosis of MS [193].

1.6.7 Prognosis

Many articles have studied the prognostic values of clinical and imaging parameters measured at study onset in predicting short-term and long-term MS evolution. A poor prognosis is generally associated with old age at MS onset, male sex, the involvement of multiple systems, a great number of relapses, and a rapid increase in EDSS score during the first few years [39] [50] [187] [188]. Not only the relapse frequency but also the type of relapse may be important. Patients with high rates of motor and sphincter relapses have an increased risk of developing a secondary progressive disease after a relapsing-remitting evolution [16].

Optic neuritis is the first sign in some MS patients [54]. Optic neuritis is the inflammation and loss of the myelin covering the optic nerve and retina. MRI measures during optic neuritis relapse give important clinical information on the risk of future MS [195] [219]. Besides, pyramidal and cerebellar scores are significantly associated with the time elapsed from disease onset to EDSS 3, EDSS 6, and the transition to secondary progression [44]

Number of relapses	Lesions	Additional criteria
two or more	Objective clinical evidence of two or more lesions or objective clinical evidence of one lesion with reasonable historical evidence of a prior relapse	None. Clinical evidence alone will suffice; additional evidence desirable but must be consistent with MS
two or more	Objective clinical evidence of one lesion	Dissemination in space, demonstrated by - \geq one T2 lesion in at least two MS typical CNS regions (periventricular, juxtacortical, infratentorial, spinal cord); OR - Await further clinical relapse implicating a different CNS site
one	Objective clinical evidence of two or more lesions	Dissemination over time, demonstrated by - Simultaneous asymptomatic contrast-enhancing and non-enhancing lesions at any time; OR - A new T2 and/or contrast-enhancing lesions(s) on follow-up MRI, irrespective of its timing; OR - Await a second clinical relapse
1	Objective clinical evidence of one lesion	Dissemination in space, demonstrated by - \geq one T2 lesion in at least two MS typical CNS regions (periventricular, juxtacortical, infratentorial, spinal cord); OR - Await further clinical relapse implicating a different CNS site AND Dissemination over time, demonstrated by - Simultaneous asymptomatic contrast-enhancing and non-enhancing lesions at any time; OR - A new T2 and/or contrast-enhancing lesions(s) on follow-up MRI, irrespective of its timing; OR - Await a second clinical relapse
0 (progression from onset)		One year of disease progression (retrospective or prospective) AND at least two out of three criteria: - Dissemination in space in the brain based on \geq one T2 lesion in periventricular, juxtacortical or infratentorial regions; - Dissemination in space in the spinal cord based on \geq two T2 lesions; OR - Positive CSF

Table 1.4: 2010 revised McDonald Diagnostic Criteria for MS.

[226]. some variables such as age, sex, and symptoms are strong predictors of shift from the disease onset to an irreversible disease phase (EDSS 4, 6, or 7); these variables became non-significant for predicting subsequent disease progress after EDSS score 4 has been reached [40].

As for imaging parameters, lesion load has been found to be an important prognostic measure in CIS patients [25] [3] [67]. A high baseline brain volume predicts better long-term clinical outcomes in RR patients whereas high baseline and greater early increase in EDSS score is significantly associated with worse outcomes [221]. In addition to the conventional imaging parameters, the advanced imaging parameters such as fractional anisotropy (FA) measurement or mean diffusivity (MD) in Diffusion Tensor Imaging (DTI) are reported as highly predictive indicators of clinical outcome [96] [56]. Another study showed that FA and MD measurements were not significantly different between different clinical subtypes but they are highly associated with the disease evolution in RR and SP patients [33].

1.6.8 Disease-modifying therapies

Although on the grounds of non-curative approaches, a number of disease-modifying therapies (DMTs) became available during the past 20 years, especially for patients experiencing relapses. These treatments reduce the relapse rate and the number of new MRI lesions. However, as the drugs act in the inflammation, rather than in the neurodegenerative mechanism of the disease, they do not stop disease progression in long-term [106] [100] [160] [73] [138]. Thus, today's DMT are partially efficient in CIS, RR and SP patients, but not in PP patients [86] [88]. The treatments in MS have two aspects: immunomodulatory drugs to reduce autoimmune disorders and drugs to relieve partially or completely worsening symptoms.

The classical immunomodulators of MS are interferon beta ($IFN\beta$) and Glatiramer Acetate (GA): two were approved for the treatment of RR and CIS patients. $IFN\beta$ has many actions such as inhibition of T-cell activation and reduction of blood-brain barrier permeability to inflammatory cells. GA acts also on the immune system by interfering with antigen presentation and induction of suppressor cells.

The effects of $IFN\beta$ and GA proved to reduce disability progression in, about 0-80% of new T2 enhancing lesions, almost 30% of relapses [159] [35]. Moreover, the use of these treatments over a long period by RR patients delayed the time to secondary progression [69] [216].

Second-line therapies such as Alemtuzumab, Natalizumab, and Mitoxantrone are also available for more progressive patients when $IFN\beta$ and GA are not efficient. Mitoxantrone is an immunosuppressive agent whose effect was proven in RR and SP patients; however, its toxicity limits its use as basic treatment [141] [28]. Alemtuzumab is also effective, even in progressive MS including SP patients where the frequency of active lesions, brain and spinal cord atrophy are elevated [36] [157]. The American Academy of Neurology in 2008 showed that Alemtuzumab was more efficient than Interferon in a 3 year-follow-up, it led to 0.77 reductions in EDSS and also to a reduction of relapse rate and disability progression.

Chapter 2

Comparison of classification methods for tissue outcome after ischemic stroke

2.1 Introduction

The five classification methods (SVM, ANN, RF, ADA, and LR) that were presented in Section 1.1 is used to predict the infarction risk in patients with stroke. The prediction of high infarction risk would help to determine the patients who are eligible for tPA treatment.

In this chapter, imaging data that has been used was voxel-based. Voxel represents a tiny cube of brain tissue in a three dimensional imaging. Each voxel can represent a million or so brain cells. We used eight MRI parameters (i.e., T2 FLAIR, ADC, CBV, CBF, TTP, MTT, TMAX, and DWI) to characterize each voxel's risk of infarction. The five classification methods predicted the infarction risk of each voxel in patients with stroke.

Let x_{ij} be the voxel of individual i observed with MRI parameter j . LR estimates the infarction risk (π_i) of voxel i by the logit function defined in following equation

$$\pi_i = \frac{1}{1 + e^{-(\beta_0 + \sum_{j=1}^8 \beta_j \times x_{ij})}} \quad (2.1)$$

where β_0 is the intercept and β_j the effect size of one unit change of MRI parameter x_{ij} on the output.

SVM separates voxels into healthy or infarcted using a linear border. The signed distance to the border can be used as the infarction risk. A positive distance represents high risk of infarction, and a negative distance low risk of infarction.

ANN involves several types of layers: an input layer composed of all data x_{ij} , an output layer that gives the final outcome, and one (or more) hidden layers between the input and the output layer. The input layer sends first the information to the next layer with an initial weight and this weight is updated after the response of the network has reached the output layer. A weighted sum (sum_i) of the responses of the neurons for voxel i is

computed defined as $sum_i = \sum_{i=1}^n x_{ij} \times w_{ij}$. The risk of infarction is obtained by applying a sigmoid function to this weighted sum:

$$f(x_{ij}) = \frac{1}{1 + \exp(-sum_i)} \quad (2.2)$$

RF builds several bootstrap samples from the original data and fits a decision tree to each sample to classify the voxel into healthy or infarcted. The risk of infarction of a voxel is computed as the percentage of trees that classified this voxel as infarcted.

ADA corrects sequentially the classification weighting the misclassified voxels with a set of decision trees h_t . According to the misclassification error of the decision tree and the weights given to the voxels, the performance α_t of each tree is computed. For a given voxel, the infarction risk is the result of the classification of each tree weighted by its performance $\sum_t^T \alpha_t \times h_t$.

The predicted risk of infarction of each voxel was dichotomized after by a threshold. A voxel with an infarction risk above the threshold was considered as infarcted. The threshold was defined as the value that minimizes the difference between observed and predicted infarcted volume.

FLAIR sequence at one month after stroke onset was used to identify the voxels corresponding to contour the boundaries of the irreversibly damaged brain tissue. The voxels in boundary were labelled as infarcted and other voxels were labelled as healthy. The predictions of classification methods were compared with the voxels observed on FLAIR sequence.

The ischemic data is highly imbalanced because the infarcted voxels are much less numerous than healthy voxels. For this reason, AUC_{prc} was used to assess the performance of the five classification methods. In addition to AUC_{prc} , AUC_{roc} , sensitivity, and specificity were calculated to compare our results with previous studies.

ADA, ANN, LR, and RF gave significantly higher AUC_{roc} values in comparison with SVM. However, there was no significant difference regarding sensitivity, and specificity, and AUC_{prc} .

67% of patients received treatment tPA after admission. Figure 2.1 shows that median of AUC_{prc} was always higher for untreated patients than treated patients for each method. However, the interquartile range of AUC_{prc} values were close for treated and untreated patients. Moreover, there was no significantly difference between AUC_{prc} values of treated and untreated patients for all classification methods (Wilcoxon rank sum test : $p\text{-value}_{LR}=0.71$, $p\text{-value}_{SVM}=0.65$, $p\text{-value}_{ANN}=0.47$, $p\text{-value}_{RF}=0.53$, $p\text{-value}_{ADA}=0.56$.)

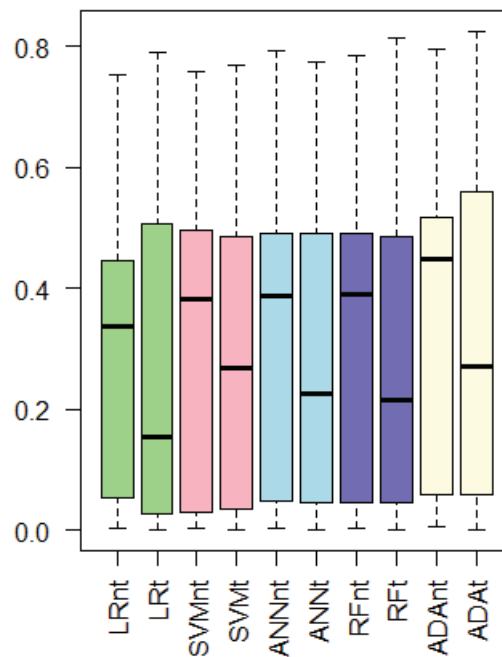


Figure 2.1: The AUC_{prc} values calculated separately for treated and untreated patients with tPA by each classification methods. The AUC_{prc} values performed for treated patients indicated adding “t” at the end of the method name on the x axis..

Besides, we analyzed if there was a difference of absolute volume error ($|\text{Observed Volume} - \text{Predicted Volume}|$) between treated and untreated patients with tPA. No signifi-

cant difference of absolute volume error was observed between treated and untreated patients (Wilcoxon rank sum test: $p\text{-value}_{LR} = 0.47$, $p\text{-value}_{SVM} = 0.99$, $p\text{-value}_{ANN} = 0.81$, $p\text{-value}_{RF} = 0.31$, $p\text{-value}_{ADA} = 0.58$).

The following article shows the details of implementation of the five methods in patients with stroke from the I-KNOW multicentre study (<http://www.i-knowstroke.eu>). The article is in preparation and will be submitted to the journal "Frontiers in Neurology".

Journal: Frontiers in Neurology

Comparison of classification methods for tissue outcome after ischemic stroke

Running headline: Comparison of classification methods in stroke MRI

Ceren Tozlu^{1,2,3,4}; Brice Ozenne^{5,6}; Tae-Hee Cho⁷; Norbert Nighoghossian⁷; Irene Klærke Mikkelsen⁸; Laurent Derex⁷; Marc Hermier⁷; Salvador Pedraza⁹; Jens Fiehler¹⁰; Leif Østergaard^{8,11,11}; Yves Berthezène^{7,8}; Jean-Claude Baron^{12,13}; Delphine Maucort-Boulch^{1,2,3,4}

¹ Hospices Civils de Lyon, Service de Biostatistique-Bioinformatique, F-69003 Lyon, France.

² Université de Lyon, F-69000, Lyon, France.

³ Université Lyon 1, F-69100, Villeurbanne, France.

⁴ CNRS, UMR5558, Laboratoire de Biométrie et de Biologie Évolutive, Équipe Biostatistique-Santé, F-69100, Villeurbanne, France.

⁵ Neurobiology Research Unit, Rigshospitalet, 9 Blegdamsvej, 2100 Copenhagen O, Denmark

⁶ Department of Public Health, Section of Biostatistics, University of Copenhagen, 5 Øster Farimagsgade, 1014 Copenhagen K, Denmark

⁷ Department of Stroke Medicine and Department of Neuroradiology, Université Lyon 1; CREATIS, CNRS UMR 5220-INSERM U1044, INSA-Lyon; Hospices Civils de Lyon, Lyon, France.

⁸ Center of Functionally Integrative Neuroscience, Aarhus University, Aarhus, Denmark.

⁹ Department of Radiology (IDI), Girona Biomedical Research Institute (IDIBGI), Hospital Universitari de Girona Doctor Josep Trueta, Girona, Spain.

¹⁰ Department of Diagnostic and Interventional Neuroradiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.

¹¹ Department of Neuroradiology, Aarhus University Hospital, Aarhus, Denmark

¹² Department of Clinical Neurosciences, University of Cambridge, Cambridge, United Kingdom.

¹³ INSERM U894, Hôpital Sainte-Anne, Université Paris Descartes, Sorbonne Paris Cité, France.

Corresponding author:

Pr. Delphine Maucort-Boulch

Service de Biostatistique-Bioinformatique

162 avenue Lacassagne

F-69003 – Lyon, France

Phone: (+33) 4 72 11 57 36 - Fax: (+33) 4 72 11 51 41

E-mail: delphine.maucort-boulch@chu-lyon.fr

Sources of funding that require acknowledgment: I-KNOW consortium was funded by the European Commission's Sixth Framework Programme (Grant 027294).

ABSTRACT

In acute ischemic stroke, identifying brain tissue at high risk of infarction is important for clinical decision-making. This tissue may be identified with suitable classification methods from magnetic resonance imaging (MRI) data. The aim of the present study was to assess comparatively the performance of five popular classification methods (Adaptative Boosting (ADA), Logistic Regression (LR), Artificial Neural Networks (ANN), Random Forest (RF), and Support Vector Machine (SVM)) in identifying tissue at high risk of infarction on human voxel-based brain imaging data. The classification methods were used with eight MRI parameters including diffusion-weighted imaging (DWI) and perfusion-weighted imaging (DWI) obtained in 55 patients. Sensitivity, specificity, the area under the receiver operating curve (ROC) as well as the area under the precision-recall curve criteria were used to compare the method performances. The methods performed equally in terms of sensitivity and specificity while the results of the area under the ROC were significantly better for ADA, LR, ANN and RF. However, there was no statistically significant difference between the performances of these five classification methods regarding the area under the precision-recall curve, which was the main comparison metric.

Keywords: brain ischemia, classification, diffusion-weighted imaging, Magnetic Resonance Imaging, machine learning, perfusion-weighted imaging

INTRODUCTION

Ischemic stroke is one of the major causes of death or long-term disability in most developed countries.¹ One of the current medical stroke therapies is intravenous thrombolysis which has to be administered as soon as possible after symptom onset. Besides, identifying the tissue at risk of infarction with an imaging tool would help decision-making in personalized treatment.²

Brain imaging based on Magnetic Resonance Imaging (MRI) provides important variable related to acute cerebral ischemia. Various MRI sequences have been developed to this end. Diffusion-weighted imaging (DWI) is a marker of irreversibly damaged tissue whereas perfusion-weighted imaging (PWI) characterizes the tissue's blood supply, and particularly the ischemic penumbra, which is at risk of infarction but still salvageable by early reperfusion.^{3,4} In the acute stroke setting, a marked discrepancy between the size of abnormal tissue identified by DWI vs. PWI¹, so called DWI-PWI mismatch, is often encountered, and the performance of DWI and PWI data combined in identifying the tissue at risk of infarction has been found to be superior to that of DWI data or PWI data alone.^{3,4}

Identifying the tissue at risk of infarction in each patient using MRI and specific statistical methods would help to determine the subject's most appropriate treatment.^{3,4} Many classification methods have been already used to provide the risk of infarction on voxel-based data. For example, in 2006, Wu et al.⁵ have proposed the use of generalized linear models to estimate the probability of infarction on the basis of diffusion- and perfusion-weighted imaging data from humans, but machine learning algorithms may outperform the generalized linear model in case of complex multimodal data. In 2010, Huang et al.⁶ used

Artificial Neural Networks (ANN) on animal imaging data and reported promising findings regarding predicting the outcome of ischemic tissue. In 2011, another study from the same authors⁷ found better results with Support Vector Machine (SVM) than with ANN on animal data. A recent extensive study by Bouts et al.⁸ compared five classification methods on experimental animal data; the five methods gave similar results. However, extrapolating these findings to man may not be appropriate.⁹ Thus, Winder et al.¹⁰ compared the prediction accuracy of three classification methods (nearest-neighbor, generalized linear model, and random forest RF) using human voxel-based stroke data; RF --as a machine learning algorithm-- performed significantly better than the two other methods.

The aim of the present study was to assess the performance of five popular classification methods (Adaptative Boosting – ADA, ANN, Logistic Regression – LR, RF, and SVM) in identifying the tissue at high risk of infarction from voxel-based human data.

MATERIAL AND METHODS

Patients

The study used data of patients from the I-KNOW multicentre study (<http://www.i-know-stroke.eu>) that included prospectively patients who underwent MRI at admission and follow-up to estimate voxel-based probabilistic maps of infarction risk. The inclusion criteria were: (1) National Institutes of Health Stroke Scale (NIHSS) ≥ 4 ; (2) DWI and PWI data consistent with an anterior-circulation acute ischemic stroke; (3) admission MRI carried out within 6 hours in case of intravenous recombinant tissue plasminogen activator (rt-PA) use or within 12 hours in case of conservative treatment. The I-Know study conformed with the Helsinki Declaration, the rules laid out by the Council of Europe Convention on Human rights and Biomedicine, Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, and with the legislation and regulations in Denmark, Germany, France, and Spain, respectively. The study was approved by the Aarhus, Hamburg, Lyon, and Girona hospitals respective regional ethics committees, and carried out after informed consent from the patients. Patients with lacunar or posterior circulation stroke, unknown time of stroke onset, unknown T2 FLAIR sequence measure or intracerebral haemorrhage on MRI were excluded. Overall, 55 patients were used for each classification method.

Image acquisition and processing

On admission, all patients underwent: i) a DWI MRI sequence (3 or 12 directions, repetition time >6000 ms, field of view 24 cm, matrix 128x128, slice thickness 3 or 5mm); ii) a Fluid-Attenuated-Inversion-Recovery (FLAIR: repetition time 8690 ms, echo time 109 ms, inversion time 2500 ms, flip angle 150°, field of view 21 cm, matrix 224x256, 24 sections, section thickness 5 mm, slice gap 1 mm); iii) a PWI MRI (echo time 30-50 ms, repetition time 1500ms, field of view 24 cm, matrix 128x128, 18 slices, thickness 5 mm, gap=1 mm, gadolinium contrast at 0.1 mmol/kg, intravenous injection 5 mL/s followed by 30 mL saline).

Diffusion-weighted sequence generated maps of DWI and apparent diffusion coefficient (ADC) parameters. PWI tracked the bolus of injected contrast agent to generate maps of cerebral blood volume (CBV), cerebral blood flow (CBF), and transit time map parameters such as the time to peak (TTP), the mean transit time (MTT), and the time-to-maximum (TMAX). Perfusion maps were computed by circular singular value decomposition of the tissue concentration curves with an arterial input function from the contralateral middle cerebral artery. Using a reference region from the contralateral normal white matter, temporal parameters normalized by subtracting the mean contralateral value and all further references to MTT, TMAX, and TTP refer to the relative parameters.

FLAIR imaging at one month after stroke onset was used to contour the boundaries of the irreversibly damaged brain tissue in stroke patients.

The parameter maps of each patient were normalized using the mean and the standard deviation of the contralateral tissue (cerebrospinal fluid excluded). The means were calculated on three consecutive slices and the standard deviation on the full brain volume. The parameter values were then centred and scaled to be comparable in effect size.

All eight MRI-based parameters (i.e., T2 FLAIR, ADC, CBV, CBF, TTP, MTT, TMAX, and DWI) were used to characterize each voxel's risk of infarction.

Classification methods

Four machine learning methods (SVM, ANN, RF, and ADA) and LR were used in this study. (For more details of methods and their settings, see the Appendix). These five methods allow identifying the risk of infarction of each voxel-based observation.

The LR estimated the infarction risk of a voxel as its probability of being infarcted after certain follow-up periods using the same above-cited combination of eight MRI parameters.¹¹

The SVM separated the observations into healthy or infarcted using a linear border. The closest observations to the border in each class were named support vectors. Then, the support vectors helped to choose the best border either through maximizing the distance between the border and the support vectors or through minimizing the number of

misclassified observations.¹² The signed distance to the border is related to the infarction risk. A zero distance corresponds to an observation located on the border, a positive distance to a high risk, and a negative distance to a low risk.

The ANN constitute a mathematical representation of natural neural networks.¹³ Each network is composed of several types of layers: (1) an input layer composed of all data, (2) an output layer giving the final outcome, and (3) one (or more) hidden layers between the input and the output layer that consist(s) of a set of neurons that process the data and are connected to the input and output layers. The input layer sends first the information to the next layer with an initial weight and the weight is updated after the response of the network is reached in the output layer. The update iterations continue until there is no change. The weighted sum of the responses of the neurons in the output layer is then computed. The risk of infarction is found by applying a sigmoid function to this weighted sum.

The other classification methods, RF and ADA, are ensemble methods that combine several decision trees. RF builds several bootstrap samples from the original data and fits a decision tree to each sample to classify the observations into healthy or infarcted.^{14,15} At each observation, the risk of infarction is then computed as the percentage of trees that classify this observation as infarcted. The ADA weights the misclassified observations using a set of decision trees.¹⁵ The classification of the first tree is performed with the same weight for each observation but the weights of the misclassified observations are increased after each decision

tree classification. The performance of each tree is computed using the misclassification error of the decision tree and the weights given to the observations. At each observation, the infarction risk is the result of the classification of each tree weighted by its performance.

Statistical analysis

The performance of the five classification methods tested here were mainly compared in terms of area under the precision-recall curve (AUC_{prc}). This criterion was selected because it summarizes the predictive ability of each method over all possible thresholds allowing for infarction prevalence. The AUC_{pr} was preferred to the area under the receiver operating curve (AUC_{roc}) because the ischemic stroke data were highly imbalanced; i.e., the infarcted voxels are much less numerous than non-infarcted voxels.¹⁶ However, as many studies have already used the AUC_{roc} to compare classification methods, this criterion was also reported here to allow result comparisons.

The results from the different methods enabled classifying each voxel into healthy or infarcted using a given threshold. The threshold was defined as the value that minimizes the difference between the observed and the predicted infarcted volume; thus, a voxel with an infarction risk above the threshold was considered as infarcted. The predicted infarction volume for each patient could be calculated as the sum of the voxels classified as infarcted. Threshold-dependent criteria such as sensitivity (Se) and specificity (Sp) were also calculated to evaluate method performances using the infarcted volume.

The leave-one-out cross-validation approach was used to avoid overfitting and provide an accurate estimation of the prediction performances of the different methods.¹⁷ Patients were removed one by one and the method applied to the remaining data. Afterwards, the prediction relative to the removed patient was performed using the same method. AUC_{prc} , AUC_{roc} , Se, and Sp were separately calculated for each of the 55 patients and the medians of these values given as evaluation criteria.

Method performances calculated with AUC_{pr} , AUC_{roc} , Se, and Sp were compared with the Kruskal-Wallis test. A higher value of AUC_{pr} , AUC_{roc} , Se, and Sp obtained by a classification method indicates a better prediction performance. The independence between treatment and gender, having hypertension, and cigarette smoking before the stroke was tested with Chi-Squared test. The age between treated and untreated patients was compared with Wilcoxon rank sum and signed rank test. The significance level set at $p < 0.05$. The free software R (<https://www.r-project.org>) was used for all statistical analyses and graphs.

RESULTS

Patients

Table 1 shows the patients' characteristics. We observed that 37 out of the 55 patients were treated with intravenous thrombolysis. The numbers of women and men were similar among treated patients (49% vs 51%) whereas there were more men than women among untreated patients (22% vs 78%). The majority of patients (65%) had hypertension before the stroke. The gender and receiving treatment were significantly independent (p -value=0.11). 30% of the patients had smoked regularly before stroke. More untreated patient has been observed smoking than treated patient (21% vs 50%) but there was no significant relationship between smoking and the treatment (p -value=0.25). In terms of infarcted volume, the median of final infarction volume was higher than the DWI lesion volume. But, we observed no significant change in infarcted volume at one month for treated and untreated patients (Paired Wilcoxon test: $p - value_{treated} = 0.08, p - value_{untreated} = 0.26$).

	Treated (n=37)	Untreated (n=18)	All patients (n=55)
Female	49%	22%	40%
Prior hypertension	62%	72%	65%
Smoking	21%	50%	30%
Age (yrs)	70 [65; 77]	67.5 [60; 74,75]	68 [62; 77]
DWI lesion volume, mL	12 [4; 27]	4 [3; 21]	11 [4; 26]
Final infarct volume, mL	12 [3; 43]	8 [2, 24]	10 [3; 40]

Table 1 – Patients' clinical characteristics at hospital admission and one month later. The patients are divided in two groups such as treated with rt-PA or untreated with rt-PA. Values are presented as median [first quartile; third quartile] – rt-PA: recombinant tissue plasminogen activator.

Classification

Table 2 presents the AUC_{prc} , AUC_{roc} , Se, and Sp values relative to the five classification methods. The AUC_{prc} values ranged between 0.20 and 0.34; the highest value was obtained with SVM and the lowest with LR. However, there was no statistically significant difference between AUC_{prc} values (p-value = 0.75). All AUC_{roc} values were higher than 0.7. RF, ADA, ANN, and LR performed the same and significantly better than SVM when AUC_{roc} values compared with Kruskal-Wallis test (p-value <0.05). The Sensitivity values were all less than 0.5

whereas the Specificity values were all larger than 0.9. There was no statistically significant difference between methods in terms of Se or Sp.

Classification method	AUCprc	AUCroc	Sensitivity	Specificity
ADA	0.29 [0.05; 0.53]	0.88 [0.83; 0.94]	0.49 [0.29; 0.65]	0.97 [0.96; 0.98]
ANN	0.28 [0.04; 0.49]	0.87 [0.82; 0.94]	0.47 [0.29; 0.64]	0.97 [0.96; 0.98]
LR	0.20 [0.03; 0.45]	0.88 [0.83; 0.92]	0.47 [0.23; 0.62]	0.97 [0.95; 0.98]
RF	0.30 [0.04; 0.48]	0.86 [0.79; 0.92]	0.47 [0.28; 0.67]	0.98 [0.96; 0.98]
SVM	0.34 [0.03; 0.48]	0.77 [0.66; 0.84]	0.48 [0.30; 0.67]	0.97 [0.96; 0.98]

Values are expressed as median [first quartile; third quartile].

Table 2 – Performance criteria of the five studied classification methods.

Figure 1 shows the AUC_{prc} and AUC_{roc} values performed by LR and SVM on each patient data. These two methods were selected because they gave respectively the highest and lowest AUC_{prc} values (Table 2). The graphs show that both AUC_{prc} and AUC_{roc} values with LR are higher than SVM values on almost all patient data.

In the case of volumes <50 mL, both AUC_{prc} and AUC_{roc} values were highly variable whereas AUC_{roc} values tended to have higher values than AUC_{prc} values. In volumes > 100 mL, AUC_{prc} and AUC_{roc} values were similar.

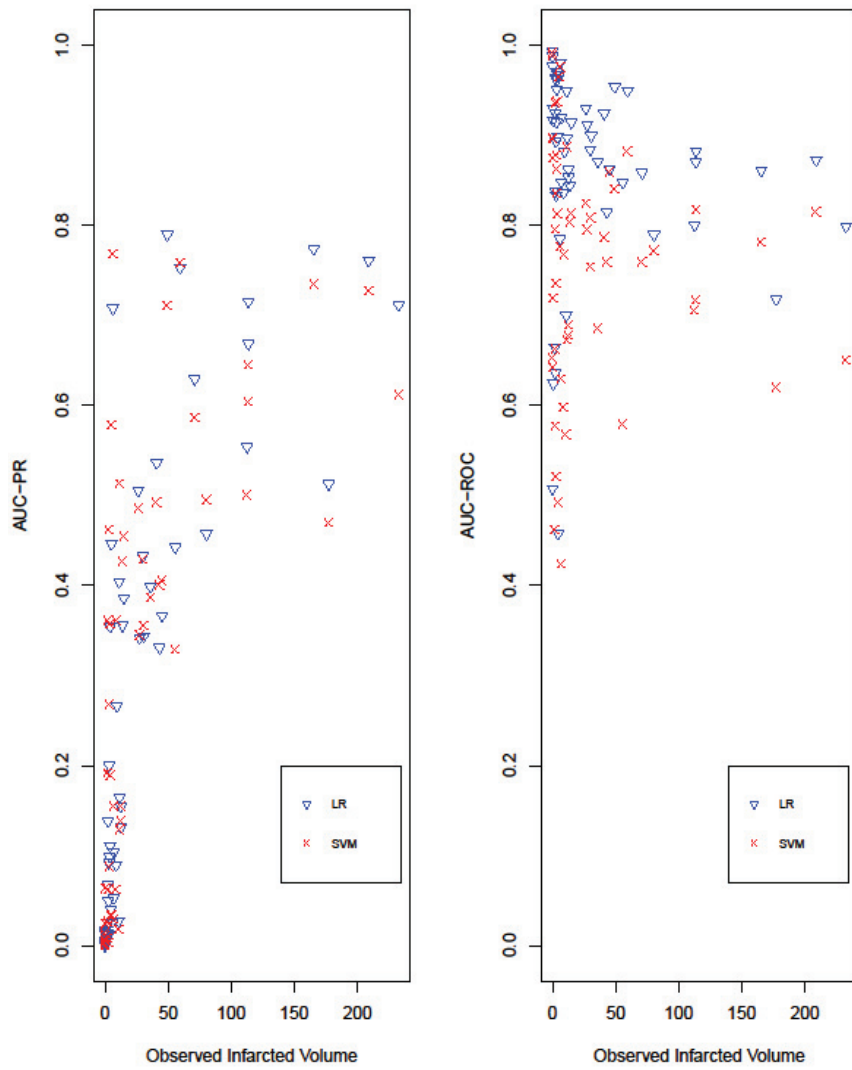


Figure 1 – Observed infarcted volumes in 55 patients at one month after stroke as expressed by the area under precision-recall curve AUC_{prc} (left panel) and the area under the receiver operating curve AUC_{roc} (right panel) obtained with logistic regression (LR) and Support Vector Machine (SVM).

Figure 2 shows the voxels of predicted infarction on illustrative examples of brain MRI slices. The rows show findings for two representative patients with respectively Regularly and irregularly-shaped infarction. ADA, ANN, LR, and RF performed approximately the same in the two patients. For the patient shown in the top row, ADA, ANN, LR, and RF overestimated infarct size, with numerous false positive voxels. In the bottom row patient, all methods, particularly SVM, underestimated final infarct size.

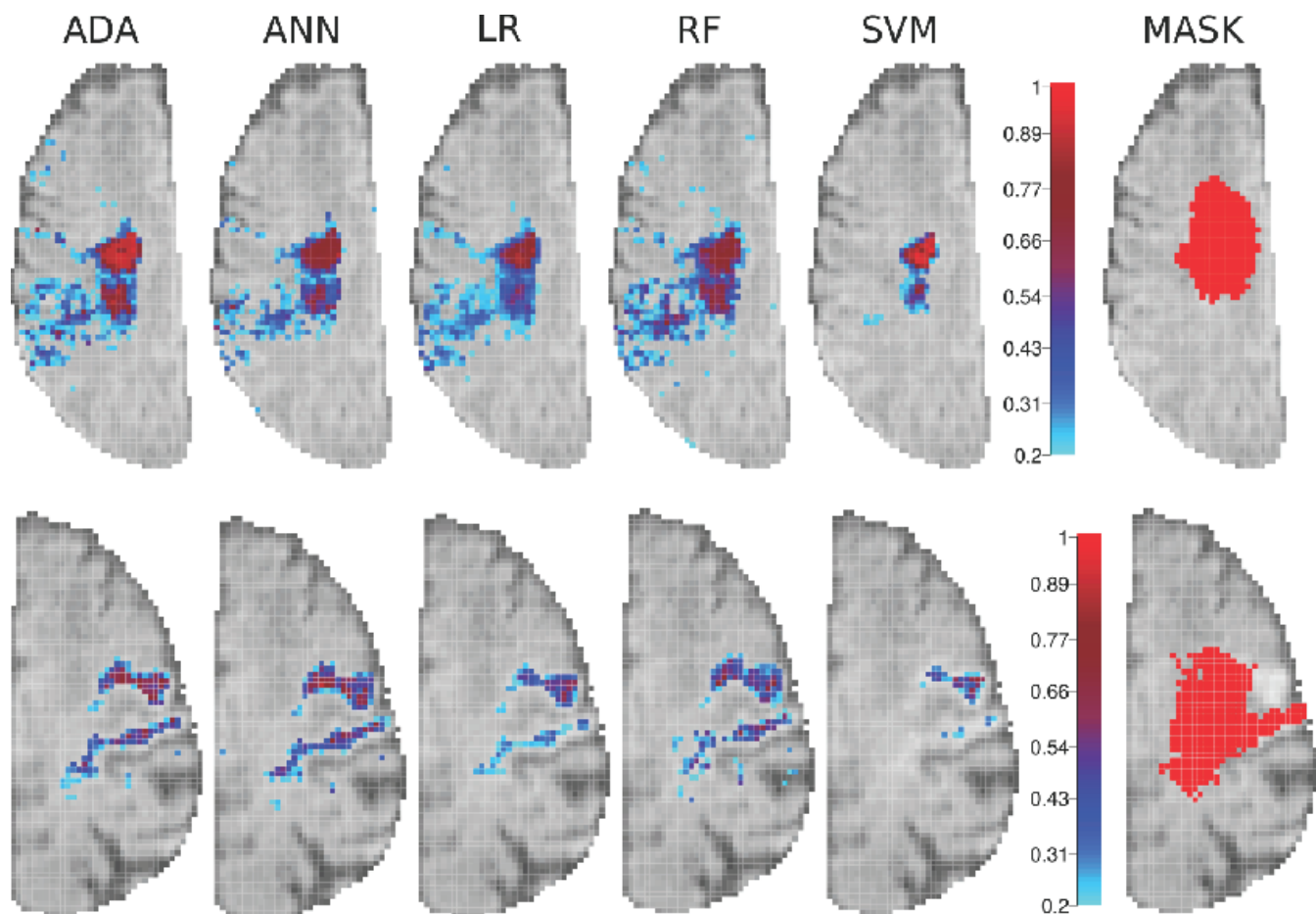


Figure 2 – Voxel-wise infarction risk prediction ([0; 1] probability) obtained with the five methods (left), and final infarct (red) as delineated on one-month post-stroke FLAIR on the right hand side, for two representative patients. Each column shows the results yielded by each method, projected onto the template MRI mask. The risk scale ranges from light blue (risk = 0) to deep red (risk = 1). ANA: Adaptive Boosting – ANN: Artificial Neural Networks – LR: Logistic Regression - RF: Random Forest – SVM: Support Vector Machine.

Figure 3 shows the bar plots of the predicted and observed infarction volumes one month after stroke for three selected patients. These patients were chosen because they showed discrepant results with SVM. The SVM method underestimated final infarct volume in patient (A), overestimated it in patient (B) and performed similarly to the other methods in patient (C). However, ADA, ANN, LR, and RF performed equally well in all three patients regardless of infarct volume.

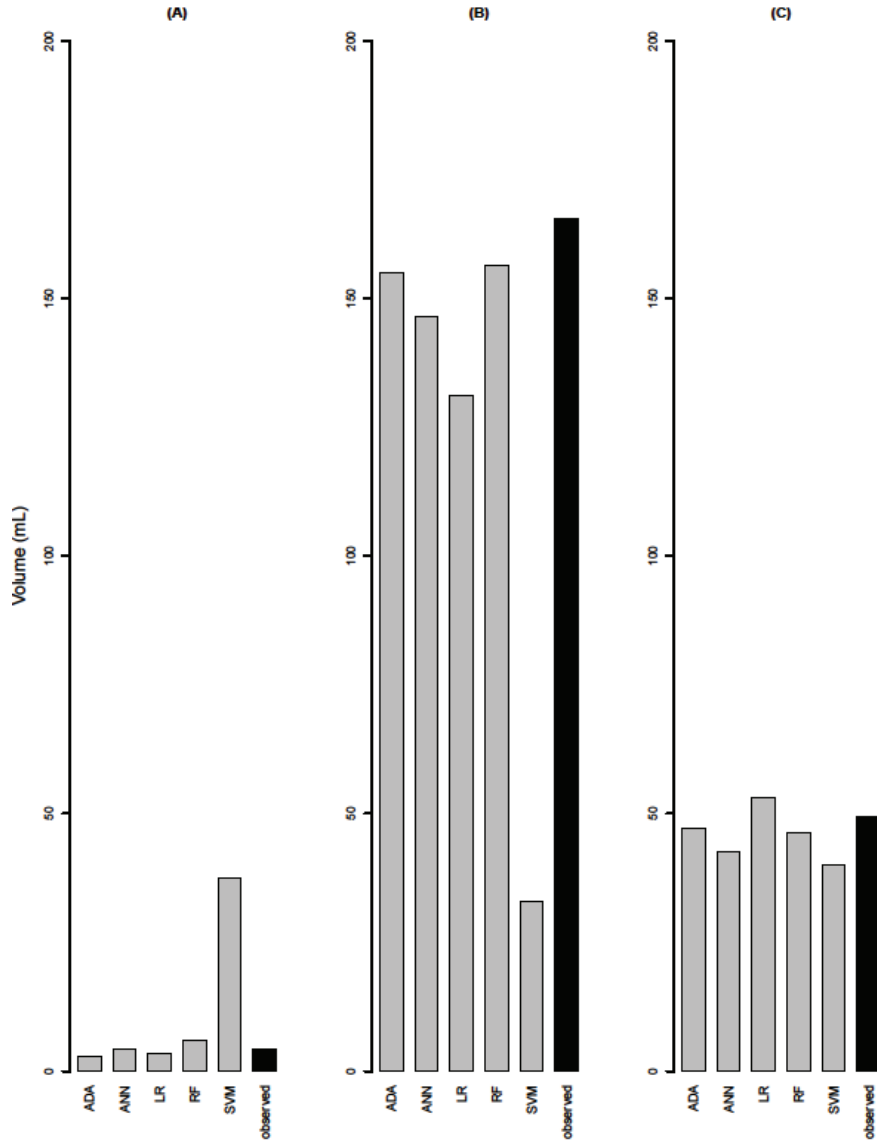


Figure 3 – Bar plots of the predicted (gray bars) and observed infarcted volumes (black bars) one month after stroke in three representative patients. ADA: Adaptive Boosting – ANN: Artificial Neural Networks – LR: Logistic Regression - RF: Random Forest – SVM: Support Vector Machine.

DISCUSSION

To our knowledge, our study is the largest to compare classification methods (especially including four machine learning methods) on imaging data from stroke patients. In this study, five classification methods to identify the brain tissue at risk of infarction were compared using voxel-based multimodal MRI data. Another large study in animals with method comparison was previously reported by Bouts et al.⁸, but was not applied onto human data. Our findings suggest no significant difference in performance of the five classification methods in terms of identification of the tissue at risk of infarction on human imaging data.

Our findings are consistent with previous results obtained based on animal data.^{7,8} However, contrary to our study, Winder et al.¹⁰ found that RF performed better than ANN and GLM. This discrepancy may be due to the use by these authors of a different performance criterion, namely Dice coefficient. Dice coefficient considers the true positives and is used to evaluate the accuracy of the predicted infarction volume for a given threshold. However, the comparison metric used in our study (AUC_{prc}) summarizes the predictive ability over all possible thresholds.

In the present study, the performance of each method was first evaluated by AUC_{prc} values, which allowed us to summarize the identification ability of the method over all possible thresholds. On the basis of this criterion, the five classification methods performed equally

well in identifying tissue at risk of infarction, in agreement with another study that used the same criterion but on animal data.⁸

Regarding the AUC_{roc} criterion, there was no significant difference in performance between ADA, ANN, LR, and RF methods. However, all performed significantly better than SVM. The latter finding may be due to the fact that SVM is essentially a binary classification method that requires an additional step able to provide the infarction risk. This gives less accurate risk predictions thus lower AUC_{roc} values.¹⁸ These results contrast with those obtained on experimental data by Huang et al.⁷ (who showed that SVM outperformed ANN) and with those obtained by Bouts et al.⁸ who showed that all methods performed equally well.

In the present study, AUC_{roc} values were always higher than AUC_{prc} values with all methods. This is due to the fact that AUC_{roc} is not sensitive to the imbalance between healthy and infarcted voxels and overestimate the infarcted volume in studies performed on data with low prevalence.¹⁶ When we compared the sensitivities and specificities of the methods, the median of Se with each method ranged between 0.4 and 0.5 while the median of Sp was over 0.9. Thus, all methods performed better in identifying healthy tissue than the tissue at risk of infarction.

Most previous studies used mainly the AUC_{prc} , the AUC_{roc} , and also the Dice coefficient to compare classification methods on human or experimental ischemic voxel-based

data.^{7,8,10} For this reason, we used the more often used criteria in our study: the AUC_{prc} and the AUC_{roc} . Regarding the AUC_{prc} , all previous studies concluded that there were no significant differences between methods. However, RF and SVM performed better than ANN regarding Dice coefficient and AUC_{roc} criteria in some studies whereas Bouts et al. concluded to similar performances of the methods they used.

The stroke patient data used in this study had a wide range of final infarction volumes. Therefore, the predicted infarcted volumes at one month obtained with different methods were quite different. Despite the heterogeneity of infarcted volumes, ADA, ANN, LR, and RF performed equally and provided homogeneous criteria values whereas SVM gave more heterogeneous values depending on the observed infarction volume.

In terms of computer resources, LR was the fastest method; with one central processing unit (2.7 GHz and 26 Go RAM), the results were obtained in 30 seconds for all 55 patients whereas SVM required approximately 30 hours. Another advantage of LR vs. other methods is that this method does not need a step of parameter selection to fit the model.

In conclusion, the five classification methods performed equally well in terms of identifying the volume of brain tissue at risk of infarction. The ADA, ANN, LR, and RF methods showed equally good performances in term of infarcted volume prediction. The results show that statistical models based on multiparametric MRI can provide valuable prognostic

information in acute ischemic stroke, with the potential to guide physicians in their time-critical decision process regarding choice of therapy.

Acknowledgements

The authors thank Jean Iwaz (Hospices Civils de Lyon) for the revision of the final drafts of the article.

Author Contribution statement

CT carried out the statistical analyses and drafted the article.

BO organized, checked, and managed the data, helped with the statistical analyses, and commented and reviewed the article.

LO carried out data collection.

D M-B organised the study and commented and reviewed the article.

Disclosure/Conflict of Interest

The authors have no specific conflicts of interest in relation with the present article.

Supplementary material

Supplementary material for this paper is available at:

<http://jcbfm.sagepub.com/content/by/supplemental-data>.

REFERENCES

1. Srinivasan A, Goyal M, Al Azri F, et al. State-of-the-Art Imaging of Acute Stroke. *RadioGraphics* 2006; 26 Suppl 1: S75-96.
2. Hacke W, Kaste M, Bluhmki E, et al. Thrombolysis with Alteplase 3 to 4.5 Hours after Acute Ischemic Stroke. *N Engl J Med* 2008; 359: 1317-1329. doi: 10.1056/NEJMoa0804656.
3. Sorensen AG, Copen WA, Østergaard L, et al. Hyperacute Stroke: Simultaneous Measurement of Relative Cerebral Blood Volume, Relative Cerebral Blood Flow, and Mean Tissue Transit Time. *Radiology* 1999; 210: 519-527.
4. Wu O, Koroshetz WJ, Østergaard L, et al. Predicting Tissue Outcome in Acute Human Cerebral Ischemia Using Combined Diffusion-and Perfusion-Weighted MR Imaging. *Stroke* 2001; 32: 933-942.
5. Wu O, Christensen S, Hjort N, et al. Characterizing physiological heterogeneity of infarction risk in acute human ischaemic stroke using MRI. *Brain* 2006; 129: 2384-2393.
6. Huang S, Shen Q, and Duong TQ. Artificial neural network prediction of ischemic tissue fate in acute stroke imaging. *J Cereb Blood Flow Metab* 2010; 30: 1661-1670.
7. Huang S, Shen Q, and Duong TQ. Quantitative prediction of acute ischemic tissue fate using support vector machine. *Brain Res* 2011: 1405; 77-84.
8. Bouts MJ, Tiebosch IA, Van Der Toorn A, et al. Early identification of potentially salvageable tissue with MRI- based predictive algorithms after experimental ischemic

- stroke. *J Cereb Blood Flow Metab* 2013; 33: 1075-1082.
9. Casals JB, Pieri NCG, Feitosa MLT, et al. The use of animal models for stroke research: A review. *Comp Med* 2011; 61: 305-313.
 10. Winder AJ, Siemonsen S, Flottmann F, et al. Comparison of classification methods for voxel-based prediction of acute ischemic stroke outcome following intra-arterial intervention. *Proc SPIE* 2017; 10134.
 11. Hosmer DW and Lemeshow S. *Applied Logistic Regression*. 2nd ed. New York: Wiley, 2000.
 12. Hsu C, Chang C, and Lin C. *A Practical Guide to Support Vector Classification*. <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>. (2003, last update 2016, accessed 19/05/ 2016)
 13. Mohiuddin KM, Mao J, and Jain AK. Artificial neural networks: A tutorial. *Computer* 1996; 29: 31-44. DOI: 10.1109/2.485891
 14. Breiman L. Random Forests. *Mach Learn* 2001; 45: 5-32.
 15. Freund Y and Schapire RE. A Short Introduction to Boosting. *Journal of Japanese Society for Artificial Intelligence* 1999; 14: 771-780.
 16. Ozenne B, Subtil F, Maucourt-Boulch D. The precision-recall curve overcame the optimism of the receiver operating characteristic curve in rare diseases. *J Clin Epidemiol* 2015; 68: 855-859.
 17. Elisseeff A and Pontil M. Leave-one-out error and stability of learning algorithms with applications. In: Suykens J, Horvath G, Basu S, Micchelli C, Vanderwalle J (eds.). *Advances in Learning Theory: Methods, Models and Applications*. NATO Science Series III: Computers

and Systems Sciences. Amsterdam: IOS Press, 2003, vol. 190, pp 11-130. (2002, accessed 19/06/2016)

18. Bishop CM. *Pattern Recognition and Machine Learning*. 1st ed. Springer. New York: Springer, 2007.

APPENDIX

We explain below the ways the methods used to calculate the infarction risk of a given n observation.

The LR estimates the infarction risk of an observation as its probability π_i of being infarcted at one month. Probability π_i is calculated using MRI parameters x_{ij} with the following logit function: $\text{logit}(\pi_i) = \alpha_{LR} + \sum_{j=1}^p \beta_{j,LR} \times x_{ij}$, where α_{LR} is the intercept and $\beta_{j,LR}$ the multiplicative risk factor.

The SVM separates the observations into healthy or infarcted using a linear border. The closest observations to the border in each class are called “support vectors”. The support vectors will help afterwards choosing the best border by maximizing the distance between the border and the support vectors and minimizing the number of misclassified observations. To find a linear border, SVM projects all observations into a higher dimensional space via a kernel function. The SVM optimizes the solution by minimizing $\|w\|^2 + C \sum \xi_i$ respecting $y_i(\langle \vec{w}, x_i \rangle + \alpha) \geq 1$, where \vec{w} is the normal (perpendicular) vector to the border and C the trade-off constant. The signed distance to the border can be used as the infarction risk. A zero distance represent a voxel that belongs to the border line, a positive distance represents a high risk of infarction, and a negative distance a low risk of infarction.

The ANN is a mathematical representation of natural neural networks. Each network involves several types of layers: an input layer composed of all data x_i , an output layer that gives the final outcome y , and one (or more) hidden layers between the input and the output layer that consist(s) of a set of neurons that process the data and are connected to the other layers. The input layer sends first the information to the next layer with an initial weight and this weight is updated after the response of the network has reached the output layer. The update iterations continue until there is no further change. A weighted sum of the responses of the neurons is then computed as: $sum_i = w_0 + \sum_{j=1}^p x_{ij} \times w_{ij}$, where w_0 is the intercept, x_{ij} the responses of the neurons of the previous layer, and w_{ij} the final updated weights. The risk of infarction is obtained by applying a sigmoid function to this weighted sum: $f(x_i) = \frac{1}{1+\exp(-sum_i)}$.

The RF builds several bootstrap samples from the original data and fits a decision tree to each sample to classify the observations into healthy or infarcted. For each split of a tree classification, the best parameter is chosen within a sample of parameters. This sample is created randomly and its size is fixed a priori (mtry). For a given observation, the risk of infarction is then computed as the percentage of trees that classified this observation as infarcted.

The ADA uses another strategy which corrects sequentially the classification weighting the misclassified observations with a set of decision trees h_t . The classification of the first tree is

performed with the same weight for all observations but the weights of misclassified observations are increased after each decision tree classification. According to the misclassification error of the decision tree and the weights given to the observations, the performance α_t of each tree is computed. For a given observation, the infarction risk is the result of the classification of each tree weighted by its performance $\sum_{t=1}^T \alpha_t \times h_t$.

Settings

SVM, ANN, RF, and ADA require parameter settings before the fit of the method.

A Gaussian Radial Basis kernel function $K(\vec{x}_1, \vec{x}_2) = \exp(-\gamma \|\vec{x}_1 - \vec{x}_2\|^2)$ was used for SVM where γ is a supplementary kernel parameter and \vec{x}_1, \vec{x}_2 are observations. To optimize the kernel parameter γ as well as the trade-off constant C, a grid-searching approach based on the misclassification error rate was used where $C \in [1, 10]$ was incremented by steps of 1 and $\gamma \in (0.001, 0.005, 0.01, 0.05, 0.1, 1)$.

For the ANN, a single hidden layer was fitted, and a grid-searching was used to optimize the number of nodes in interval [14, 25] with steps of 1 and a decay parameter $\in (0.001, 0.005, 0.01, 0.05, 0.1, 0.5$ and 1). The decay parameter can be seen as a regularization parameter that avoids over-fitting.

For the RF, the number of parameters used at each node to perform the classification was set at $m_{rty} = 3$, the number of trees optimized in the interval [1, 100] by increments of 1, and the performances of the combinations compared using the misclassification error rate.

For the ADA, the optimal number of trees was searched in the interval [1,200] by increments of 1 and the performances of the combinations compared using also the misclassification error rate.

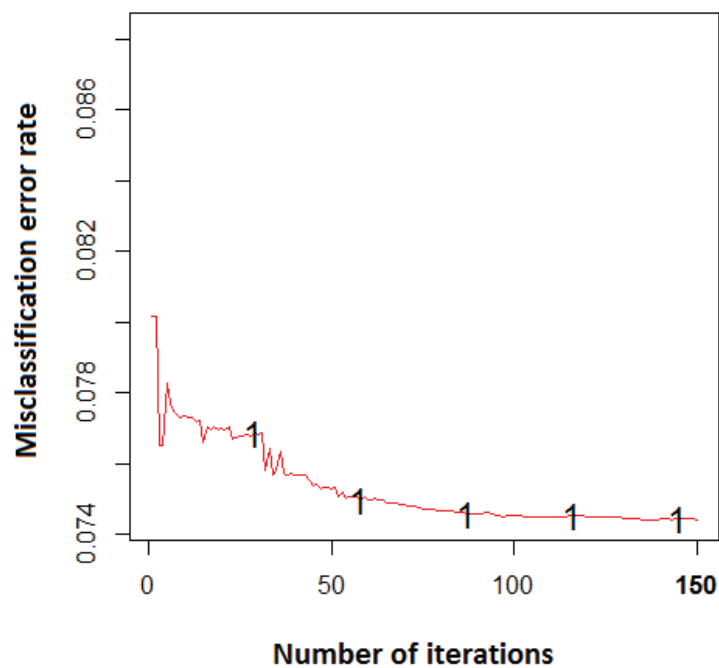
With ADA, RF, ANN, and LR, the threshold that minimizes the difference between the observed and the predicted infarcted volume was set between 0 and 1 with steps of 0.001. With SVM, this threshold was set between -3 and 3 with steps of 0.001.

Supplementary figures

Parameter optimization was made using a grid-searching approach that provides the misclassification error rate with ADA, ANN, RF, and SVM algorithms. The parameters which gave the lowest misclassification error rate were chosen.

With ADA, the optimal number of iterations (which is also the number of trees) is searched in the interval [0, 150] with increments of 1. Supplementary figure 1 shows that the misclassification error calculated with ADA is nearly the same when the number of trees is

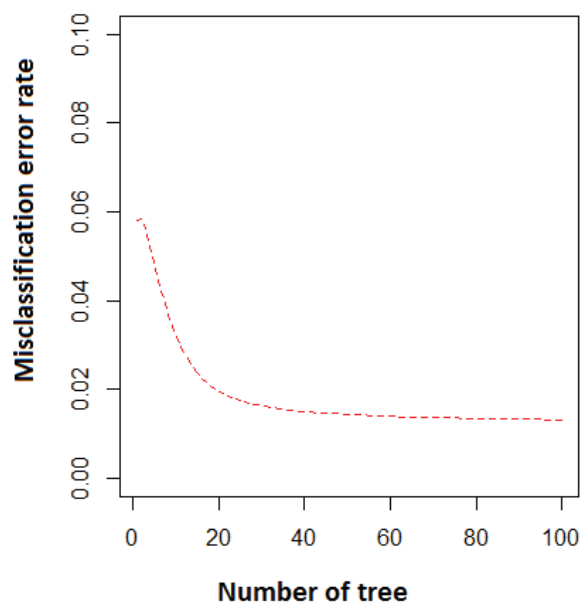
between 100 and 150 and that it reached the minimum at 150 trees when the misclassification error rate was almost $\xi = 0.075$.



Supplementary figure 1 – Misclassification error rate according to the number of iterations in adaptive boosting analysis.

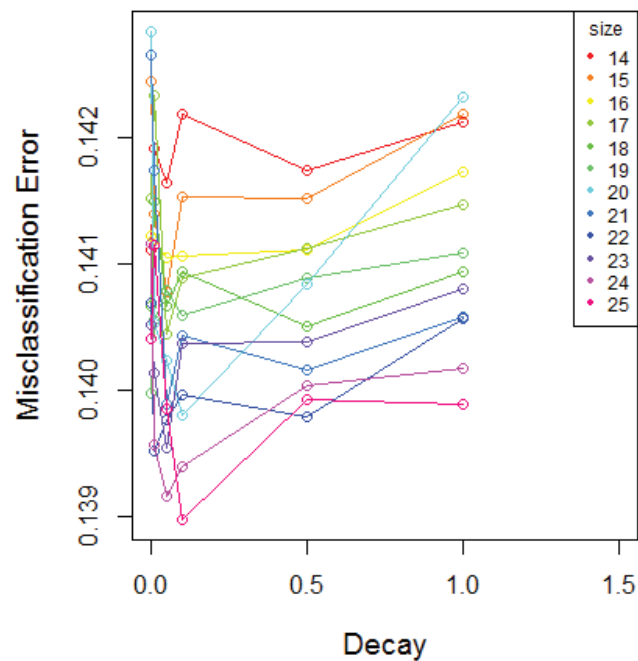
With RF, the number of variables at each node (mtry) was fixed at 3 as suggested by Breiman and the optimal number of trees (ntree) searched from 1 to 100 with increments of 1 (Freund

&, Schapire. *J Japanese Soc Artif Intell* 1999;14:771-780) The misclassification error rate on the Y-axis in Supplementary figure 2 is almost constant for a number of trees ranging between 20 and 100 and the misclassification error rate reached the minimum value at ntree=100. So, the couple of parameters (ntree, mtry) = (100, 3) was used; the misclassification error dropped down to nearly $\xi = 0.05$.



Supplementary figure 2 – Misclassification error rate according to the number of trees in the random forest analysis.

With ANN, the optimization required finding the number of units in the hidden layer. This was searched in the interval [14, 25] with increments of 1 and the optimal decay was searched among values 0.001, 0.01, 0.05, 0.1, 0.5, and 1. Supplementary figure 3 shows the changes of the misclassification error in function of the decay value, the number of units being indicated by different colours. The minimum misclassification error rate ($\xi = 0.139$) was obtained with the parameter couple (size, decay) = (25, 0.1).



Supplementary figure 3 – Misclassification error rate according to the number of neurons (size) and the decay in artificial neural network analysis.

Chapter 3

Clustering of individual disability progression trajectories of multiple sclerosis patients

Contents

3.1	Introduction	118
3.2	Materials	120
3.2.1	Clinical data	120
3.2.2	Image Acquisition and Processing	120
3.3	Methods	121
3.4	Application of clustering methods	123
3.4.1	Statistical analyses	124
3.5	Results	126
3.5.1	Descriptive results	126
3.5.2	Cluster analysis results	134
3.6	Discussion	145

3.1 Introduction

Multiple sclerosis (MS) is the most frequent disabling neurological disease in young adults. While its etiology remains unknown, MS is a demyelinating, inflammatory, and chronic disease of the central nervous system. The evolution of the disease and the risk of developing permanent disability are very different from one patient to another [80]. MS patients are actually classified into four major forms based on the clinical evolution, measured by the Expanded Disability Status Scale (EDSS). Several patients ($\approx 85\%$) were diagnosed earlier as a clinically isolated syndrome (CIS) during the clinical exam. CIS patients present then relapsing-remitting (RR) evolution. RR can shift afterward to secondary-progressive (SP) with or without superimposed relapses [124]. 15% of MS patients present a primary progressive (PP), which is characterized by a continuously worsening symptoms without relapses since diagnosis [129].

Besides clinical examination, Magnetic Resonance Imaging (MRI) helps diagnosing and monitoring disease evolution. Conventional MRI, including T1-weighted and T2-weighted imaging, has an important specificity and sensitivity in detecting pathological tissue damage and providing valuable predictive information on the evolution of the disease [62] [67]. However, advanced MRI such as Diffusion Tensor Imaging (DTI) and Spectroscopic Imaging (MRSI) have higher specificities in detecting microstructural damages in white matter than conventional imaging [56] [76] because they provide complementary information on diffusion alteration and metabolism [62] [63] [61]. DTI is more able to distinguish MS subtypes as patients with different subtypes present different diffusivity patterns [56]. Therefore, we propose using DTI data together with conventional imaging in different clinical forms of MS. The FA parameter has been chosen among DTI measurements because it is more sensitive in detecting microscopic changes related to inflammation [85].

MS affects especially the white matter of the brain and the lesions appear mostly around the ventricles and periventricular white matter. The floor of the lateral ventricles forms a region called Corpus Callosum (CC). So, CC is the one of the regions the most frequently

affected by lesions in MS [14] [77] [74]. In 93% of MS patients, lesions were found in CC [77]. Moreover, it is the major myelinated bundle of the brain and, provides the connection between cortical and subcortical regions between two brain hemispheres. Therefore, the effects of lesions in CC are more severe compared to the lobar white matter. For this reason, this work used imaging parameters measured in CC.

Several studies demonstrated that callosal changes measured with DTI were correlated with cognitive and physical disability [198] [127] [173] [242]. FA measurements were significantly lower in rostrum, body, and splenium part of CC versus a control group [87] [181] [233]. The association between EDSS scores and CC atrophy as measured with conventional MR was reported in some studies [242] [192]. However, no significant correlation was found between disability and callosal atrophy as measured by conventional MR in RR patients [12]. This can be due to the insufficiency of conventional imaging in CC, thus, DTI parameters measured in CC can be more suitable for MS patients. Therefore, the FA parameter measured in CC was used in this study.

In previous studies, classification of MS patients into two clinical subtypes was performed with machine learning methods based on conventional and advanced imaging data at a given time point [215] [115] [97] [150]. However, clinical subtypes were also identified currently using EDSS score change over time. The originality of the present study is to cluster the patients on the basis of clinical and imaging longitudinal data (not at a given time point). The clusters might represent distinct evolution forms of MS patients. First, a clustering method was performed on EDSS score trajectories to compare the obtained clusters with actual clinical subtypes. Second, we clustered the MS patients using joint EDSS score and imaging trajectories. For this reason, the joint trajectories obtained with EDSS score and FA values were used to find out whether conventional imaging parameters could improve cluster results. Besides, one advanced imaging parameter was chosen regarding its correlation with EDSS score. The less correlated advanced imaging parameter was used to establish the joint trajectories. Finally, the results obtained with single and joint trajectories were compared to determine whether conventional and advanced

parameters could improve clustering.

Moreover, cluster analyses were performed only in patients with relapses during disease progression. Thus, PP patients were excluded from the data set and the cluster analyses were performed only on CIS, RR, and SP patients.

3.2 Materials

3.2.1 Clinical data

Eighty patients fulfilling the Mac Donald criteria were included in a standardized clinical and MRI protocol within the frame of the AMSEP project in Lyon Neurological Hospital. This population was divided into four groups of MS patients depending on the clinical form: CIS (n=12), RR (n=27), SP (n=16), and PP (n=25). The patients were followed up with standardized clinical and MRI examination every six months during the first three years then at one-year intervals during two years. Clinical examination included age, disease duration since onset, the Expanded Disability Status Scale (EDSS), and the Multiple Sclerosis Functional Composite (MSFC) with its three dimensions: Timed 25 Foot Walk (T25-FW), 9-Hole Peg Test (9-HPT), and Paced Auditory Serial Addition Test (PASAT). Five patients were excluded from the cohort because one or more MSFC component(s) could not be measured at study onset.

3.2.2 Image Acquisition and Processing

MS patients underwent an MR examination on a 1.5T Siemens Sonata system (Siemens Medical Solution, Erlangen, Germany) using an 8-channel head-coil. The MR protocol consisted in the acquisition of a sagittal 3D-T1 sequence ($1 \times 1 \times 1 \text{ mm}^3$, TE/TR = 4/2000 ms) and an axial 2D-spin-echo DTI sequence (TE/TR = 86/6900 ms; 2×24 directions of gradient diffusion; $b = 1000 \text{ s.mm}^{-2}$, spatial resolution of $2.5 \times 2.5 \times 2.5 \text{ mm}^3$) oriented

in the AC-PC plane.

The lesion load was measured with FLAIR sequence and the white and gray matter volumes were measured with T1 imaging without gadolinium contrast agent. Among four DTI parameters, the fractional anisotropy (FA), measured in CC was used in this study.

3.3 Methods

Packages developed by Genolini et al. **kml** and **kml3d** [78] were used to cluster, respectively, single and joint trajectories according to EDSS score and imaging parameters.

kml and **kml3d** packages allow clustering the trajectories using k-means algorithm. K-means is based on the Expectation-Maximization (EM) method. First, the number of clusters is set. Second, k-means is run to assign randomly the trajectories to clusters. Third, the mean trajectory is calculated for each cluster. The mean of each cluster is called the cluster center. The last step calculates distances between the trajectories and the cluster centers and assigns each trajectories to the nearest clusters.

kml The **kml** package is used to cluster single trajectories. Let S a set of n trajectories and y_{itX} be the longitudinal outcome measured for individual i regarding variable X at different time points t . So, $y_{i.X} = (y_{i1X}, y_{i2X}, \dots, y_{itX})$ is called a single trajectory (or single variable trajectory) for subject i . In **kml** package, distance metric by default is the Euclidian distance. The Euclidian distance between two single trajectories was calculated as below:

$$Dist(y_i, y_j) = \sqrt{\frac{1}{t} \sum_{k=1}^t (y_{ik} - y_{jk})^2} \quad (3.1)$$

where y_i and y_j are the single trajectories of subject i and j ; t represents time.

kml3d The kml3d package clusters joint trajectories established with more than one variables. Let $y_{i.A}, y_{i.B}, \dots, y_{i.M}$ be the single trajectories relative to variables A, B, \dots , M of subject i. We define the joint trajectory as:

$$y_{i..} = \begin{pmatrix} y_{i.A} \\ y_{i.B} \\ \vdots \\ y_{i.M} \end{pmatrix} = \begin{pmatrix} y_{i1A} & y_{i2A} & \dots & y_{itA} \\ y_{i1B} & y_{i2B} & \dots & y_{itB} \\ \vdots & \vdots & \ddots & \vdots \\ y_{i1M} & y_{i2M} & \dots & y_{itM} \end{pmatrix} \quad (3.2)$$

where each line corresponds to the single trajectory of each variable and t represents time.

Let $y_{1..}$ and $y_{2..}$ be the joint trajectories of two individuals. The distance between these joint trajectories is defined as $d(y_{1..}, y_{2..})$ and it can be considered as the distance between two matrices. One of the two ways to calculate $d(y_{1..}, y_{2..})$ is to compute first t distances between m rows of each column, where m is the number of variables. So, t distances $d_t(y_{1t.}, y_{2t.})$ are obtained. Then, these t distances are combined using a distance function. The second way to calculate $d(y_{1..}, y_{2..})$ is to calculate first the distance between m lines, where m is the number of variables. Therefore, m distances $d_X(y_{1.X}, y_{2.X})$ are computed. Finally, m distances are combined by a distance function to provide the distance between two joint trajectories.

The default distance between two joint trajectories in kml3d packages is the Euclidean distance which is calculated as:

$$Dist(y_{i..}, y_{j..}) = \sqrt{\sum_t \sum_X (y_{itX}, y_{jtX})^2} \quad (3.3)$$

where t and X represent respectively the time and the variable.

Quality criteria Packages kml and kml3d allow five different non-parametric criteria: 1)Calinski & Harabatz, 2)Calinski & Harabatz with Kryszczuk variant, 3)Calinski &

Harabatz with Genolini variant, 4)Ray & Turi and 5)Davies & Bouldin criterion [78] [29] [117] [170].

The criterion value for the k th cluster was calculated as:

$$C(k) = \frac{\text{Trace}(B)}{\text{Trace}(W)} \times \frac{n - k}{k - 1}$$

where $\text{Trace}(B)$ is the between-cluster covariance matrix, $\text{Trace}(W)$ is the within-cluster covariance matrix, n is the number of the trajectories and k is the number of clusters.

A quality criterion was used to help choosing the number of clusters k that maximizes the criterion $C(k)$ was used to choose the best number of clusters.

3.4 Application of clustering methods

Cluster analyses were performed first with all patients to obtain a general appropriate clustering for all clinical subtypes. Then, for the purpose of studying only the patients who had relapses during disease progression, PP patients were excluded from data and the analyses were performed only with CIS, RR and SP patients.

kml and **kml3d** packages were respectively used to cluster first according to EDSS score alone the according to EDSS score & imaging data. Before implementation, these packages request

- the interval of cluster number
- a quality criterion
- the distance metric
- the assignment of the trajectories to clusters

- number of iterations performed by algorithm

The optimal number of clusters was searched in the interval [2; 6]. The Calinski & Harabatz criterion (defined in equation 1.30) was chosen as this was the default criterion. *randomALL* method was used to assign randomly each trajectory to a cluster at the beginning of the clustering analyses. The Euclidian metric was used to calculate the distance between trajectories because it is used by default in *kml* and *kml3d* packages. The k-means algorithm was repeated 100 times with each number of clusters.

Cluster analyses were performed first on single EDSS score trajectories. Second, one of the conventional imaging parameters was used jointly with the EDSS score to examine whether a conventional imaging parameter could enhance the clustering performed on single EDSS score trajectories. A correlation test was performed between EDSS score and conventional imaging parameters to choose the best conventional imaging parameter. The less correlated parameter was used to build the joint trajectories. Third; as a advanced imaging parameter, the FA parameter was used jointly with the EDSS score to find out whether a advanced imaging parameter could improve the clustering. Afterwards, the clustering results obtained with single and joint trajectories were compared with the Calinski & Harabatz criterion.

The imaging parameters were normalized as defined in the Equation 1.28 before the clustering of the joint trajectories.

3.4.1 Statistical analyses

Because most data did not follow the normal distribution, medians, 1st, and 3rd quartiles were used to describe the data. Consequently, non-parametric statistical tests were used to find out differences in clinical and imaging parameters among the four clinical subtypes. The Kruskal-Wallis rank sum test was performed first to compare the distributions of

variable values among the four clinical subtypes. In case this test indicated that at least one median was different from the others, a Mann-Whitney U test with Bonferroni correction was performed to compare the medians two by two. Statistical significance was considered at $p < 0.05$. The statistical analyses were performed using R software with version 3.4.0 (2017-04-21). (<https://www.r-project.org/>)

3.5 Results

3.5.1 Descriptive results

Descriptive results of clinical data

Table 3.5.1 shows the main demographic and clinical characteristics of the cohort: the total number of patients, the number of women, the median and the range of disease duration, age, EDSS, pyramidal, cerebellar, and MSFC scores as well as its three components (T25FW, 9HPT and PASAT) as observed at study onset. T25FW and 9HPT scores were the means of two consecutive trials and 9HPT scores were observed with the dominant hand. The characteristics were given first for all patients and then for each clinical subtype.

Number of women of RR, PP and SP patients was higher than number of males. The disease durations at study onset were significantly different among the four clinical subtypes ($p\text{-value} < 0.05$). However, there was no significant difference between RR and PP patients' disease durations ($p\text{-value} = 0.51$). The age at study onset was also significantly different among the four clinical subtypes. Age was significantly higher in PP patients than in the other clinical subtypes ($p\text{-value} < 0.05$). However, age at study onset was similar in CIS, RR, and SP ($p\text{-value} = 0.40$).

EDSS, pyramidal, and T25FW scores at study onset were not significantly different between PP and SP patient ($p\text{-value}_{EDSS,(PP-SP)} = 0.53$, $p\text{-value}_{pyramidal,(PP-SP)} = 0.82$, $p\text{-value}_{T25FW,(PP-SP)} = 0.96$). However, EDSS, pyramidal, and T25FW scores were significantly different between other couples of clinical subtypes (i.e., CIS vs RR, CIS vs PP, CIS vs SP, RR vs PP, and RR vs PP) ($p\text{-value} < 0.05$).

Cerebellar, MSFC, 9HPT and PASAT scores were similar between CIS and RR patients, and between PP and SP patients ($p\text{-value} > 0.05$). Moreover; these four scores were significantly higher in PP and SP patients than in CIS and RR patients ($p\text{-value} < 0.05$).

	All patients n=75	CIS n=12	RR n=26	PP n=15	SP n=22
Number of women	47	6	22	10	19
Disease duration	5.93 (3.41-9.07)	0.86 (0.27- 2.27)	4.68 (2.51-7.76)	5.56 (4.67-7.75)	11.18 (8.14- 15.35)
Age	29.48(24.22 - 35.49)	29.15 (27.93 - 35.18)	26.59 (23.47 - 34.00)	35.61 (32.04 - 39.10)	32.11 (24.10 - 35.09)
EDSS	4.00 (2.00 - 4.50)	0.50 (0.00 - 1.25)	2.00 (2.00 - 4.00)	4.00 (4.00 -4.50)	4.50 (4.00 -5.00)
Pyramidal score	2.00 (0.00- 3.00)	0.00 (0.00- 0.00)	1.00 (0.00 - 2.00)	3.0 (2.0 - 3.0)	2.50 (2.00 - 3.00)
Cerebellar score	1.00 (0.00- 2.50)	0.00 (0.00- 0.00)	0.00 (0.00 - 1.00)	2.00 (1.00 - 3.00)	2.50 (1.25 - 3.00)
MSFC score	0.13 (-0.48 - 0.50)	0.68 (0.30 - 0.73)	0.45 (0.09 - 0.63)	-0.42 (-0.79 -0.03)	-0.43 (-0.88 -0.01)
Timed 25 Foot Walk	5.25 (4.42 - 6.60)	4.25 (4.02 - 4.54)	4.83 (4.25 - 6.44)	6.55 (4.95 - 7.68)	5.90 (5.31 - 7.32)
9 Hole Peg Test	22.30 (19.35 - 26.70)	18.85 (17.55 - 19.95)	20.52 (18.06 - 23.12)	24.55 (21.43 - 34.88)	27.22 (23.60 - 32.31)
PASAT	44.0 (32.0 - 54.0)	18.85 (17.55-19.95)	20.52 (18.06 -23.12)	24.55 (21.43 - 34.88)	27.22 (23.60 - 32.31)

Table 3.1: Demographic and clinical characteristics of 75 patients measured at study onset, Median (range).

The observed individual and mean EDSS trajectories that correspond to the clinical subtypes are shown in Figure 3.1. We observed that the mean EDSS trajectories were much more stable than individual EDSS trajectories. The mean trajectories of severe patients (i.e., PP and SP) were strictly above EDSS score 4 whereas the mean trajectories of CIS and RR were below EDSS score 4 which is the limit of fully ambulatory without aid.

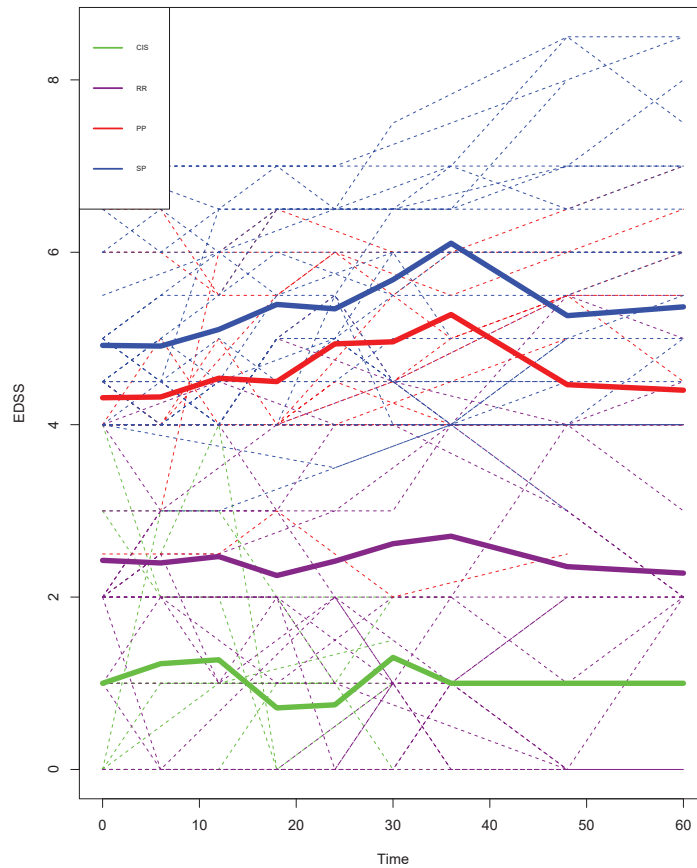


Figure 3.1: The horizontal axis shows the time in month and the vertical axis EDSS score. The dotted lines represent the individual EDSS trajectories and the solid lines the mean EDSS trajectories of each MS subtype.

Descriptive analyses of imaging data

Table 3.5.1 shows the number of imaging scans and the following conventional and advanced imaging measures: gray matter volume (GMV), white matter volume (WMV), lesion load (LL) and fractional anisotropy (FA) for each clinical subtypes. The number of imaging scans were similar in RR, PP, and SP patients (≈ 7.0 scans) (p-value > 0.05) but number was relative to CIS patients was lower (p-value < 0.05) because they were classified afterward as RR patients.

The imaging parameters were significantly different among the four clinical subtypes. However, GMV and WMV did not differ significantly between PP and SP patients (p-value > 0.05). Moreover, GMV and WMV were significantly higher in CIS and RR patients than in PP and SP patients (p-value < 0.05). The lesion load values ranged between 6 and 37 mm^3 . When the lesion load was compared between couples of clinical subtypes, there was no significant difference between RR and PP (p-value=0.86). The fractional anisotropy values ranged between 0.54 and 0.62. The highest FA values were obtained in CIS patients and were significantly different from the other subgroup' FA values (p-value < 0.05). In addition, FA values did not differ between RR and PP patients (p-values=0.38).

	All patients n=75	CIS n=12	RR n=26	PP n=15	SP n=22
Scan number	6 (6-7)	4 (3.75-5.25)	7 (6-7)	6.5 (6-8)	7 (6-8)
GMV	826.2 (801.3- 858.1)	858.3 (845.7- 883.2)	841.1 (820.7- 873.4)	811.2 (781.9- 822.9)	822.0 (791.5- 832.4)
WMV	1433 (1384-1470)	1484 (1457-1508)	1448 (1412-1477)	1420 (1346-1462)	1394 (1362-1441)
LL	15.10 (7.17- 32.97)	6.65 (4.29- 8.44)	14.37 (6.54- 24.32)	14.33 (11.00- 27.69)	36.32 (20.32- 41.52)
FA	0.59 (0.57-0.62)	0.62 (0.60-0.64)	0.60 (0.58 - 0.62)	0.59 (0.57- 0.61)	0.54 (0.48- 0.59)

Table 3.2: Conventional and advanced imaging parameters measured in 75 patients at study onset, Median (1st quartile - 3rd quartile) (GMV: Gray Matter Volume, WMV: White Matter Volume, LL: Lesion Load, FA: Fractional Anisotropy).

The observed individual and mean trajectories of imaging parameters with different colors for each clinical subtype were provided in Figure 3.2. The mean trajectories of CIS and RR patients were always above the mean trajectories of PP and SP patients regarding GMV, WMV, and FA, because, we observed higher values of GMV, WMV, and FA in the less affected MS patients. There was a slight decrease of GMV, WMV, and FA mean trajectories in RR patients. Moreover, we observed an increase of the lesion load in CIS patients after the 20th month. However, we did not observe a significant increase or decrease of the mean trajectories of any imaging parameters.

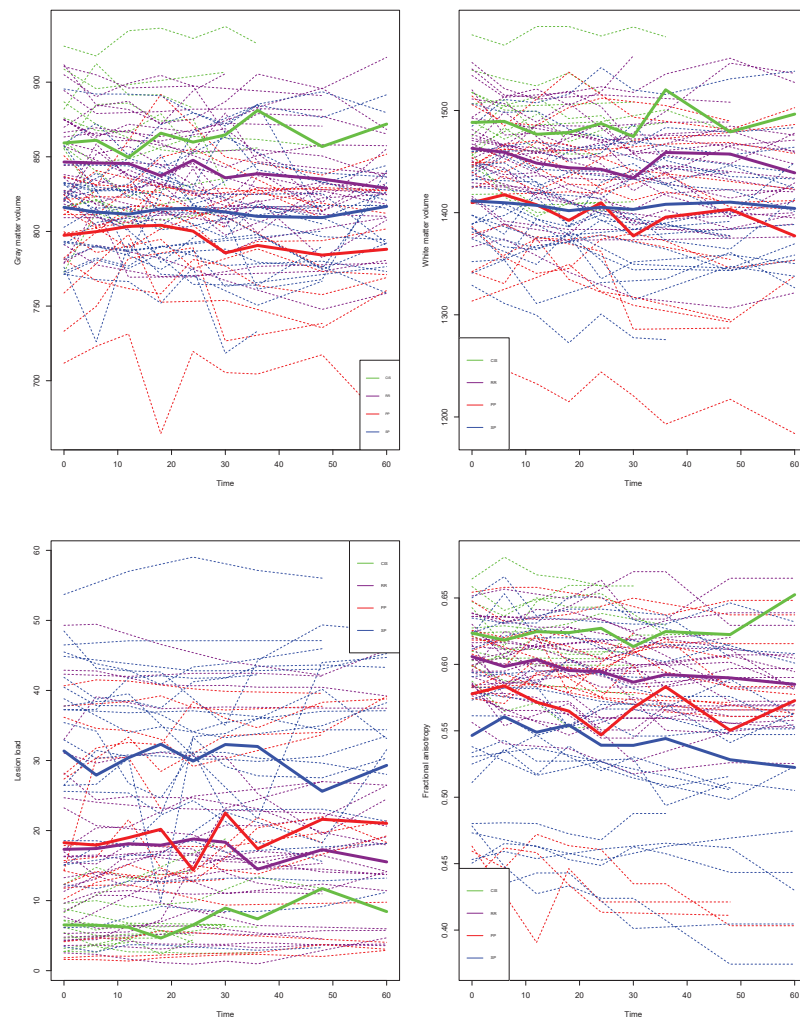


Figure 3.2: The individual (dotted line) and mean (thick solid line) of GMV, WMV, LL, and FA trajectories observed in CIS, RR, PP, and SP patients.

Regression analyses between conventional and advanced imaging parameters

Figure 3.3 shows the regression analyses on FA values (advanced imaging parameter) according to the LL, GMV, and WMV (conventional imaging parameters) at study onset. There was a negative and strong correlation between FA and LL ($\rho = -0.68$). However, there was a positive and weak correlation between FA and GMV ($\rho = 0.27$). The correlation between FA and WMV was strong ($\rho = 0.64$). As FA was influenced by axonal and myelin loss, these results showed that, in this cohort, the lesions might be mostly located in white matter.

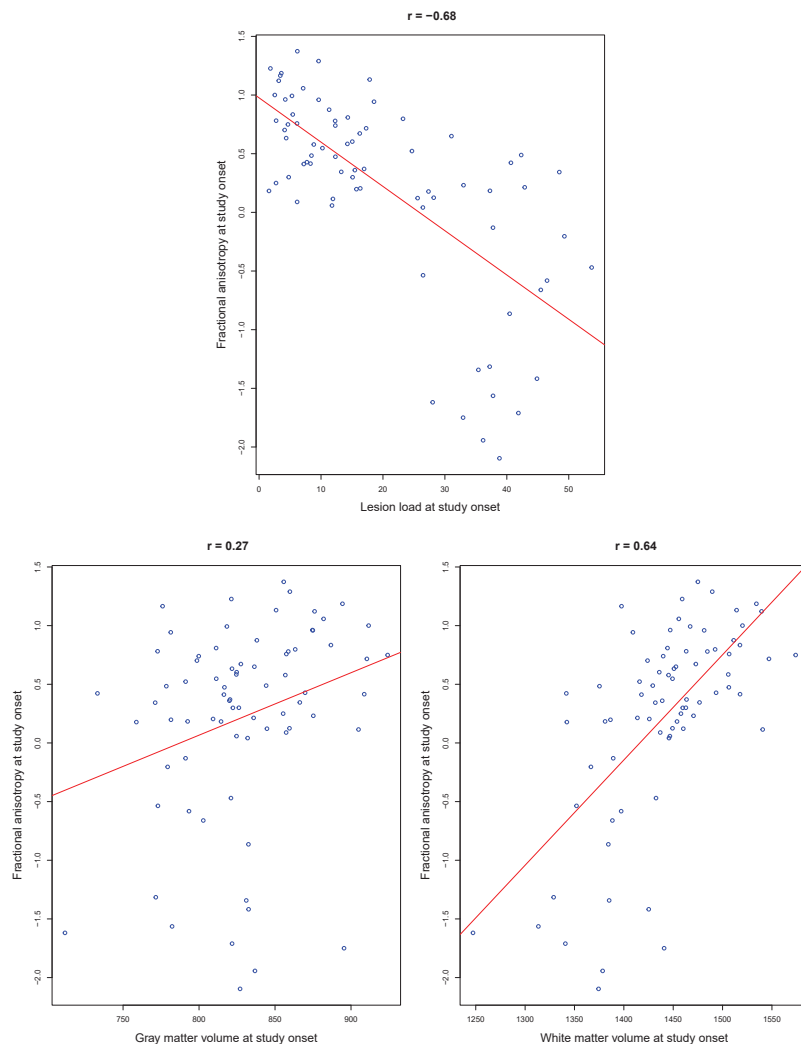


Figure 3.3: Regression analysis between conventional and advanced imaging parameters. The correlation coefficient value was indicated on the top of each graph. The red line corresponds to the linear regression line.

Regression analyses between imaging parameters and EDSS score

Figure 3.4 shows the regression analyses of imaging parameters (FA, LL, GMV, and WMV) with EDSS score at study onset. The correlation coefficient helped choosing the less correlated imaging parameters with EDSS score in order to establish the joint trajectories. All imaging parameters had weak correlation with EDSS score ($|\rho| \leq 0.55$). FA, GMV, and WMV had a negative correlation with EDSS score. This means that a decrease of FA, GMV, and WMV values causes the increase of EDSS score. However, the correlation between LL and EDSS score was positive ($\rho = 0.39$). Thus, the highest EDSS scores were observed in the patients who had highest lesion load.

The correlation coefficients were low between EDSS and LL ($|\rho| = 0.39$), and between EDSS and FA ($|\rho| = 0.42$). For this reason, LL and FA were separately used to establish the joint trajectories with EDSS score.

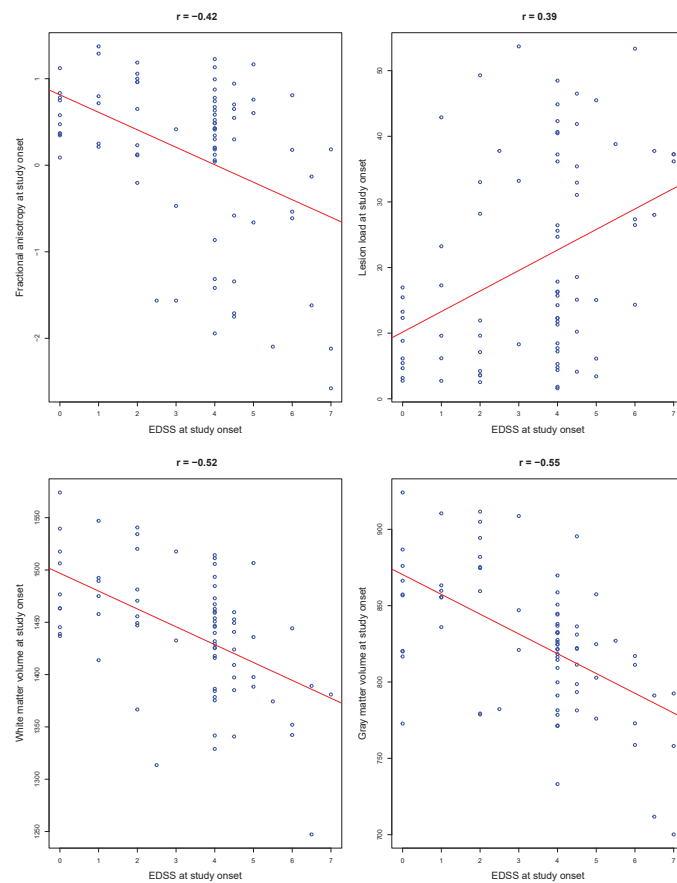


Figure 3.4: Regression analyses between EDSS and imaging parameters.

3.5.2 Cluster analysis results

In this section, we will present cluster analyses results performed on single EDSS trajectories, and joint EDSS & LL and EDSS & FA trajectories.

As indicated in section 3.4, the choice of the best number of clusters was made in the [2; 6] interval. The Calinski-Harabatz criterion was the highest with 3 clusters in both single and joint trajectories and all clinical subtypes as well as in CIS, RR, and SP patients. We present thus the results obtained with 3 clusters.

Clinical subtypes	Variables used for trajectories	2	3	4	5	6
CIS, RR, PP, and SP	EDSS	188	192	180	156	141
	EDSS & LL	84	88	77	68	62
	EDSS & FA	51	74	60	55	53
CIS, RR, and SP	EDSS	147	156	142	120	108
	EDSS & LL	72	79	62	56	51
	EDSS & FA	45	55	47	44	43

Table 3.3: Calinski-Harabatz criterion results obtained with 2, 3, 4, 5 and 6 clusters performed on single and joint trajectories in all subtypes and for CIS, RR and SP patients.

Cluster analysis results in all clinical subtypes

Figure 3.5 shows that 43% of the patients are found in cluster A. Cluster A represents the trajectory with stable evolution over time around EDSS score 4. Also, cluster A contains RR (n=12), PP (n=10), and SP (n=12) patients. However, cluster B contains CIS (n=12), RR (n=14) and PP (n=1) patients with moderate disease evolution over time. Cluster C helped identifying the more severe progressive patients: 4 PP and 10 SP patients. Unlike clusters A and B which had a stable progression, cluster C showed a mean trajectory that increases over time.

The table that corresponds to Figure 3.5 shows that CIS patients are found in one cluster (cluster B) whereas RR, PP, and SP patients are distributed over two or three clusters. Thus, we could distinguish -for example- between PP and SP patients who may have a stable evolution around EDSS score 4 or may have an aggressive evolution that increases over time above EDSS score 4.

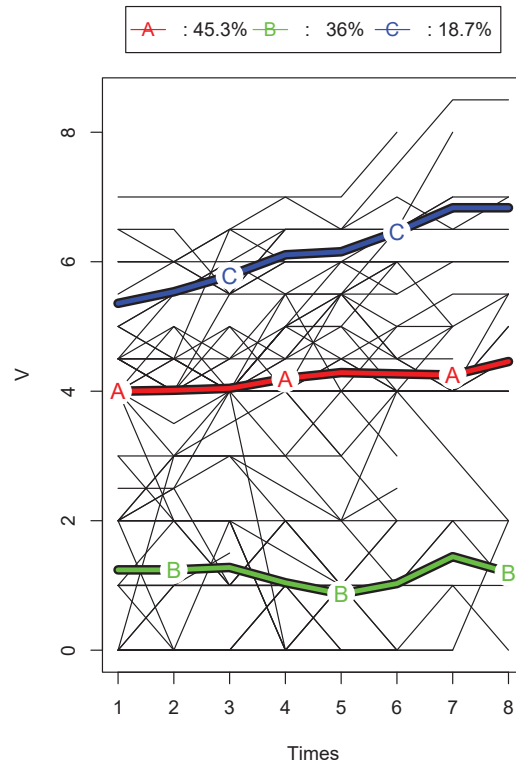


Figure 3.5: Clustering results with graphical representation and corresponding table for single EDSS trajectories of all clinical subtypes. The individual trajectories are presented with black lines and the mean trajectories of the clusters are presented with colored thick solid lines. The proportions of individuals assigned to each cluster are given at the top of the graph.

	A	B	C
CIS	0	12	0
RR	12	14	0
PP	10	1	4
SP	12	0	10

Figure 3.6 shows the graphical representation of the cluster analyses performed on respectively on the joint EDSS & LL and joint EDSS & FA trajectories. For each cluster analysis, 2 graphs are provided: one for the EDSS score and another for the imaging parameter. Figure 3.6 shows that the mean trajectories of three clusters are stable evolutions over time.

The results of the cluster analysis performed on EDSS & LL trajectories were slightly different from the results obtained with single EDSS trajectories. The proportions of the patients in the three clusters were very close (A: 34.7 %, B: 33.3 % and C: 32%). Clusters A and B included RR, PP, and SP patients. However, cluster C included only CIS and RR patients. The evolutions of EDSS trajectories of the cluster A and B were close but the EDSS evolution of cluster C was lower. Cluster B had a lower lesion load evolution than cluster A. These cluster results allowed identifying the patients with EDSS score increases despite lower lesion loads.

The advantage of joint EDSS & LL trajectories cluster was its ability to distinguish patients with different lesion load evolutions even when they have similar EDSS evolution. Clusters A and B showed similar EDSS evolutions over time. However, the lesion loads of patients in cluster A were higher than in cluster B.

Cluster results obtained with EDSS & FA trajectories showed that the proportions of patient changed slightly in comparison with the results obtained with joint EDSS & LL trajectories (A: 44 %, B: 36 % and C: 20%). Cluster B included patients with moderate evolution (12 CIS and 15 RR patients) whereas cluster C included patients with severe evolution (11 SP, 3 PP, and 1 RR patients). The graph that corresponds to the joint EDSS & FA clustering shows that cluster A had a severe EDSS evolution despite a moderate FA evolution.

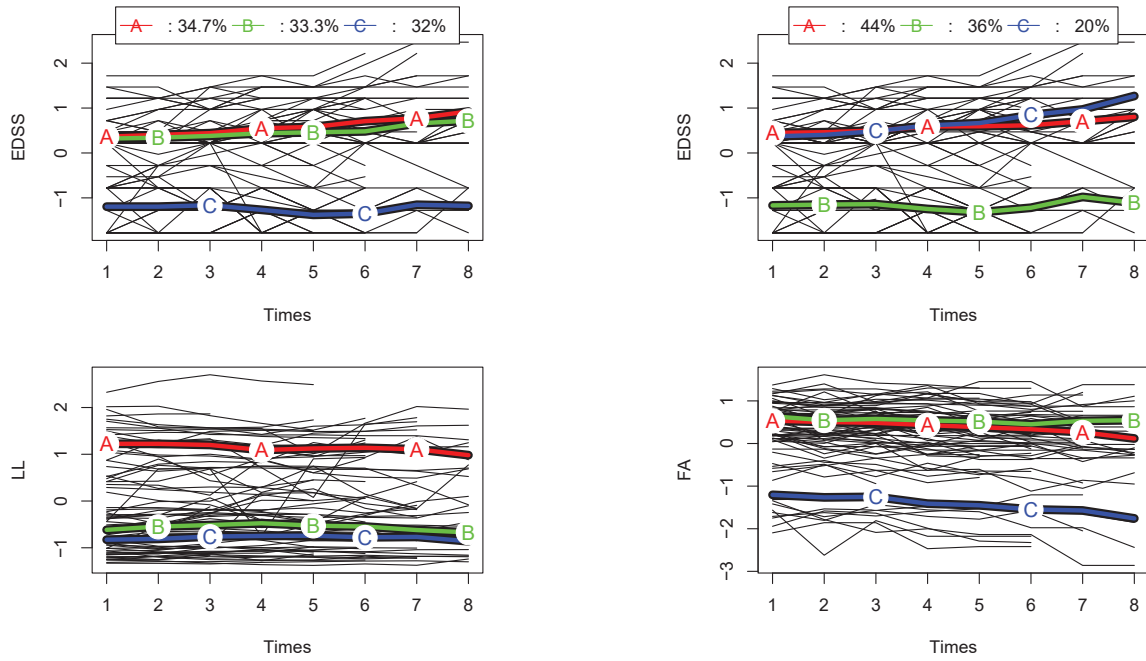


Figure 3.6: Clustering results with graphical representations for respectively EDSS & LL trajectories (left side), and EDSS & FA trajectories (right side) with corresponding clustering tables for all clinical subtypes. The individual trajectories are represented with black lines and the mean trajectories are represented with colored thick solid lines. The proportions of individuals assigned to each cluster are given at the top of the graph.

	A	B	C
CIS	0	0	12
RR	5	9	12
PP	5	10	0
SP	16	6	0

	A	B	C
CIS	0	12	0
RR	10	15	1
PP	12	0	3
SP	11	0	11

When these results were compared with the results obtained with single EDSS trajectories, the classification tables were similar but the cluster analysis performed on joint EDSS & FA trajectories helped identifying one more RR patient and one more SP patient who may have severe progressions. The observed EDSS scores of RR patient began at EDSS 2 at study onset and reached at EDSS 5.5 after 5 years. The observed EDSS score of the SP patient began at EDSS 5 at study onset and increase to EDSS 3 after 5 years. There was one less PP patient in cluster C compared to the cluster analysis performed on single EDSS trajectories. This PP patient had EDSS 6.5 at 5 years and his EDSS score did not change over time. Thus, joint EDSS & FA trajectories performed better than single EDSS trajectories in identifying the patients who had severe progressions.

When the clustering results were compared with the Calinski-Harabatz criterion, Calinski-Harabatz value of single EDSS trajectory clustering was much higher than the values obtained with joint trajectories. Moreover, the Calinski-Harabatz values of joint trajectory clustering were close ($CH_{EDSS} = 192$, $CH_{EDSS\&LL} = 88$, and $CH_{EDSS\&FA} = 74$).

Table 3.4 shows the number of patients, the median, 1st, and 3rd quartiles of clinical and imaging variables measured for each cluster obtained with the EDSS trajectory clustering. Age, disease duration, T25FW, and 9HPT, GMV, LL, and FA values at study onset in the three clusters are presented and compared. All clinical and imaging variables were significantly different among the three clusters (Kruskal-Wallis, p -value <0.05). When the variables were compared between two clusters; age, 9HPT, and 25FW values of the patients in cluster C were significantly higher than other clusters. However, the disease durations were similar between patients in clusters A and C. The GMV was significantly higher for the patients in cluster B and there was no significant difference between GMV values of cluster A and cluster C. However, the LL of the patients in cluster B was significantly lower and the LL values were similar in patients in sluster A and Cluster C. FA values were similar in clusters A and B but were significantly higher compared to the FA values in cluster C.

Variable	Class A N=9	Class B N= 63	Class C N=3
Age	40.76 (36.00, 44.56)	29.86 (26.85, 36.58)	41.84 (40.54, 44.49)
Disease Duration	7.65 (4.76, 11.19)	2.15 (0.86, 4.64)	10.18 (6.83, 15.35)
9HPT	23.40 (21.05, 24.89)	18.40 (17.48, 20.30)	34.50 (28.35, 37.39)
25FW	5.72 (5.01, 6.53)	4.25 (4.05, 4.65)	8.30 (7.06, 13.26)
GMV	821.1 (791.7, 835.5)	863.3 (845.6, 884.4)	816.5 (791.5, 826.4)
LL	17.09 (8.90, 39.65)	8.83 (5.79, 16.21)	26.91 (15.06, 36.81)
FA	0.45 (-0.13, 0.69)	0.47 (0.29, 0.89)	0.10 (-1.14, 0.52)

Table 3.4: Clinical and imaging variables according to the clusters obtained with single EDSS trajectories. Median (1st, 3rd quartile). the p -value is the result of Kruskal-Wallis test which was performed to compare the medians of variable values in the three clusters.

Cluster analysis results in CIS, RR and SP patients

Figure 3.7 is a graphical representation of the cluster analysis performed on single EDSS trajectories in CIS, RR, and SP patients. The graph shows two stable mean trajectories (clusters A and B). Cluster A was around the EDSS score 4 and cluster C progressed around the EDSS score 1. However, the mean trajectory of cluster C increased over time. The table that corresponds to Figure 3.7 shows that CIS patients were in one cluster (cluster B) whereas RR and SP patients were separated into two different clusters. Cluster B included CIS and RR patients who had moderate disease evolution. However, 10 PP patients who had severe progression over time were located in cluster C.

These cluster results were the same with cluster analyses performed on all clinical subtypes. This means that the exclusion of PP patients did not change the clustering of CIS, RR, and PP patients.

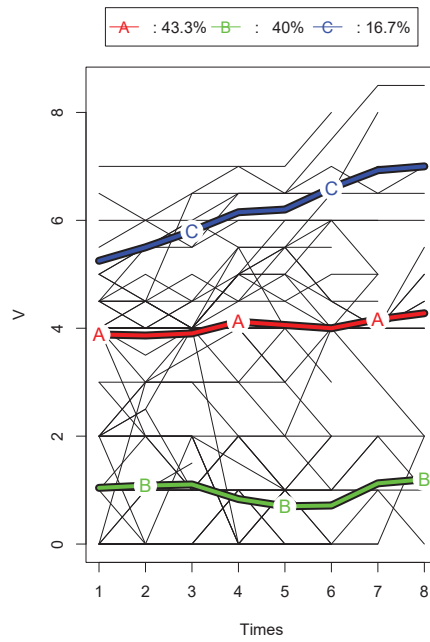


Figure 3.7: Mean trajectories of the clusters obtained with single EDSS score trajectories in CIS, RR, and SP patients.

	A	B	C
CIS	0	12	0
RR	14	12	0
SP	12	0	10

The graph on the right side of Figure 3.8 shows the mean trajectories of the clusters obtained with the joint EDSS & LL trajectories of RR, PP and SP patients. Clusters B and C had similar EDSS evolutions and they included RR and SP patients. Cluster A included CIS ($n=12$) and RR ($n=11$) patients, and the mean EDSS trajectory of cluster A was lower vs. clusters B and C. However, cluster B was quite different from cluster C in terms of lesion load. The patients in cluster B had higher lesion load. This was confirmed by the classification table: cluster B included more severe patients (SP, $n=16$) than cluster C (SP, $n=6$).

The identification of patients with high lesion load did not change when the cluster analysis

was performed excluding PP patients. We also observed that 5 RR and 16 SP patients had a higher lesion load evolution when the cluster analysis was performed on all clinical subtypes.

The graph on the right side of Figure 3.8 shows the mean EDSS & FA trajectories performed with RR, PP and SP patients. Cluster B and C had similar EDSS evolution whereas cluster A had a lower EDSS evolution over time. The clustering table showed quite similar results to those obtained with joint EDSS & LL trajectories. In these two clusterings, there were two clusters included RR and SP patients which had severe evolutions. However, the cluster with moderate EDSS evolution included CIS and RR patients.

The clustering table that corresponds to the joint EDSS & FA trajectories for CIS, RR, and SP show similar results to those obtained with all clinical subtypes. Two clusters with similar EDSS evolutions included 12 RR and 22 SP patients. This means that identifying RR and SP patients with clinically severe evolution did not change after excluding of PP patients.

When we compares the clustering results obtained with single EDSS vs. joint trajectories, the clusters were not significantly different. However, we obtained highest Calinski & Harabatz criterion value with the cluster analysis performed with single EDSS trajectories ($CH_{EDSS} = 156$). In addition, Calinski & Harabatz criterion values were much lower but similar for joint EDSS & FA and EDSS & LL trajectories ($CH_{EDSS \& LL} = 79$ and $CH_{EDSS \& FA} = 55$). The Calinski & Harabatz criterion result of the single EDSS trajectory clustering was three times higher than that obtained with joint EDSS & FA trajectory clustering. However, the number of patients in each cluster were quite similar for the cluster analysis performed with single EDSS and joint EDSS & FA trajectories. Because the patients in the clusters was different even the number of the patients were similar.

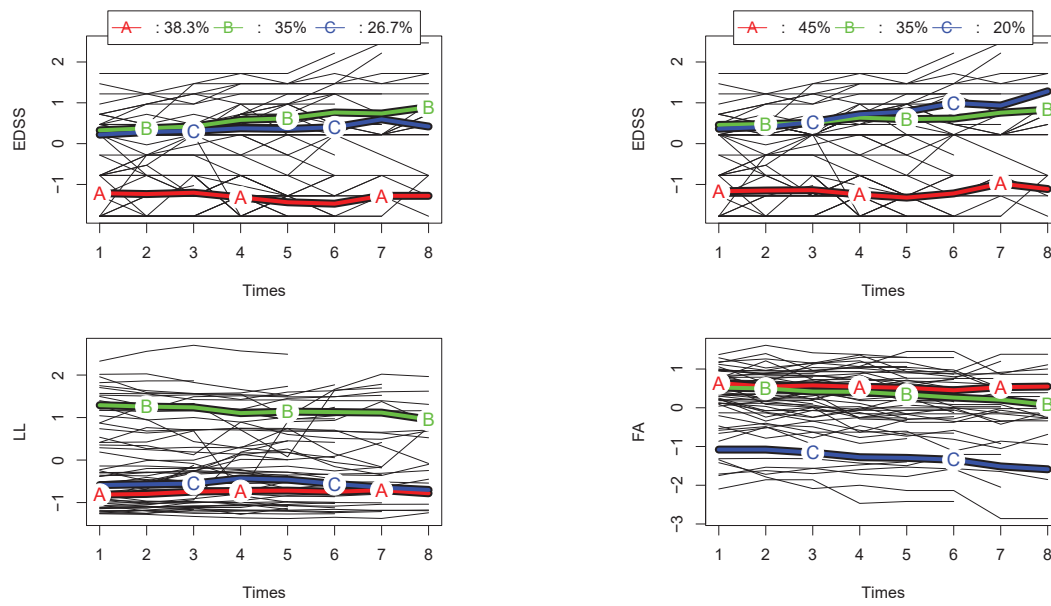


Figure 3.8: Clustering results of respectively EDSS score & LL (at left side) and EDSS score & FA (at right side) trajectories obtained with CIS, RR, and SP patients.

	A	B	C
CIS	12	0	0
RR	11	5	10
SP	0	16	6

	A	B	C
CIS	12	0	0
RR	15	10	1
SP	0	11	11

3.6 Discussion

An extension of the k-means method was performed to cluster trajectories of MS patients using clinical and imaging data. To our knowledge, this is the first study to cluster multiple sclerosis patients using longitudinal clinical and imaging data. First, the single EDSS trajectories were clustered to identify distinct forms of individual trajectories. Then, lesion load and FA variables were added one by one jointly to the EDSS trajectories to analyze their contribution to the clustering results. We showed that the EDSS trajectory clustering performed better than joint trajectory clustering. Moreover, the clustering analysis was able to distinguish patients with severe, stable and moderate evolution.

In previous studies, patients classification was performed between two clinical subtypes with machine learning methods based on conventional and advanced imaging characteristics [215] [115] [97] [150]. Some studies classified the patients with Support Vector Machine using their diffusion features [150] [115]. Taschler et al. [215] used spatially informed models and compares them with SVM method. They demonstrated that spatially informed models performed better and that MRI scans of MS lesions inform better about specific subtypes of the disease.

The classification of patients on the basis of changes over time is important because MS patients are classed regarding their clinical disability evolution over time. Moreover, the clustering analysis performed in this study helped identifying distinct evolution forms of disability trajectories. This could be useful to identify patients who may have a more important progression, so as, to change their medical treatment.

The clustering analysis performed with the single EDSS trajectories was able to identify clinical and imaging differences among the patients in the three different clusters. The clinical and imaging variables measured at study onset were significantly different among the clusters. The disease duration and LL were significantly lower in the moderate evolution cluster whereas GMV was significantly higher. Moreover, age, 9HPT, and 25FW

were significantly higher whereas FA was significantly lower in the cluster that showed severe evolution. This means that the initial values of these variables may allow predicting which patient may have a moderate or a severe disease progression.

In single and joint trajectory clustering, CIS patients were always included in one cluster that reflects at moderate evolution. However, the trajectories of RR, PP and SP patients were partitioned into at least two clusters and this may be explained by more complex disease features related to their heavy inflammation and neurodegeneration.

The clustering is actually possible with unsupervised, supervised or model-based methods. The most important advantage of the unsupervised clustering methods used in this study is that they do not require prior knowledge such as parameter estimations. They are also easy to implement, fast and efficient in terms of computational cost. However, the disadvantage of the unsupervised clustering methods is that there is no powerful quality criterion that allows choosing the number of clusters and assessing clustering accuracy. Because of this, we had to use the quality criterion proposed by default in `kml` and `kml3d` packages. Moreover, the accuracy of clustering results was tested against the observed evolution of each patient.

In our work, the additional use of imaging parameters did not significantly change the clustering results obtained with single EDSS trajectories. Moreover, the Calinski & Harabatz criterion value of single EDSS trajectory was higher than Calinski & Harabatz criterion value of joint trajectories. This may be due to the use of the lesion load in the whole brain, instead of a local load and the use of the FA in a specific part of the brain (in CC). Measuring the FA in the regions where lesions occur might be important distinguishing patient trajectories. The location and size of the lesions are also very important and play a significant role in clinical disability.

The present study shows the clustering results performed taking into account the distances between trajectories. However, these results can be extended by taking into account the variance between trajectories and the shapes of the trajectories and by introducing

demographic variables such as age and gender. This type of extension may improve the results because the disease evolution was found heterogeneous between and within clinical subtypes and because clustering may be more powerful with more variables.

Chapter 4

Modelling of individual disability
evolution of multiple sclerosis
patients using clinical and imaging
data

Modeling individual disability evolution of multiple sclerosis patients based on longitudinal multimodal imaging and clinical data with the latent class mixed model

Ceren Tozlu^{1,2,3,4}, Dominique Sappey-Marinier^{2,3,5,6}, François Cotton^{5,6}, Sandra Vukusic⁷, Françoise Durand-Dubief⁷, Delphine Maucort-Boulch^{1,2,3,4}

¹ Hospices Civils de Lyon, Service de Biostatistique-Bioinformatique, F-69003 Lyon, France.

² Université de Lyon, F-69000, Lyon, France.

³ Université Lyon 1, F-69100, Villeurbanne, France.

⁴ CNRS, UMR5558, Laboratoire de Biométrie et de Biologie Évolutive, Équipe Biostatistique-Santé, F-69100, Villeurbanne, France.

⁵ Centre de Recherche en Acquisition et Traitement de l'Image pour la Santé, (CREATIS, UMR CNRS 5220 & U1044 INSERM), Université Claude Bernard-Lyon1, Université de Lyon, France

⁶ Service de Radiologie, Centre Hospitalier Lyon-Sud, Hospices Civils de Lyon, Lyon, France

⁷ Service de Neurologie A, Hôpital Neurologique, Hospices Civils de Lyon, France

⁸ CERMEP-Imagerie du Vivant, Université de Lyon, France

Corresponding author: Delphine Maucort-Boulch

E-mail: delphine.maucort-boulch@chu.yon.fr

Phone: (+33) 4 78 86 57 96

ABSTRACT

Background: The individual disease evolution of multiple sclerosis patients is very different from one patient to another. Therefore, the prediction of the long-term disability evolution based on clinical and imaging information is actually highly difficult. Magnetic Resonance Imaging is an important tool to distinguish the healthy and abnormal brain tissue, to monitor the disease evolution, and for decision-making in personalized treatment for multiple sclerosis patients.

Objective: We aimed to develop a patient-specific model to predict the individual disease evolution of multiple sclerosis patients based on demographics, clinical and imaging information at study onset.

Methods: 75 patients (12 CIS, 26 RR, 15 PP, and 22 SP) followed within 5 years were included in our study. The latent class linear mixed model was used to consider the individual and unobserved subgroups' variability in multiple sclerosis. First, the clinical model was established with demographic and clinical variables to predict the disease evolution measured with the Expanded Disability Status Scale. Then, the multimodal imaging variables were added one by one, two by two, and three together to investigate their contribution on the clinical model. The predictive accuracies of clinical and combined models were compared according to the Bayesian Information Criterion (BIC).

Results: The clinical model gave higher BIC value than the combined models. The means of the posterior probabilities performed by the clinical and combined models were over 0.89. The clinical model was able to cluster the patients into three latent classes such as stable evolution class (n=63, 84%), severe evolution class (n=9, 12%), and moderate evolution class (n=3, 4%). The disease duration at study onset was significantly different among three latent classes.

Conclusion: The latent class linear mixed model was able to build a well-fitted predictive model for the disability evolution of multiple sclerosis patients considering the evaluation criteria.

Keywords: Multiple sclerosis, predictive modeling, latent class linear mixed model, heterogeneity, long-term disability, individual trajectory

INTRODUCTION

Multiple sclerosis (MS) is the most frequent disabling neurological disease in young adults. While its etiology remains unknown, MS is a demyelinating, inflammatory, and chronic disease of the central nervous system. The evolution of the disease and the risk of developing permanent disability are very different from one patient to another [Goldenberg, 2012]. Today's neurologists' challenge is to predict individual disability evolution on the basis of clinical, biological and imaging data.

MS patients show highly different evolution profiles and this evolution profile is not unique per patient and can change over time. MS patients are currently classified into four major subtypes, on the basis of clinical evolution as measured by the Expanded Disability Status Scale (EDSS) score. The four clinical subtypes are qualified as: the clinically isolated syndrome (CIS), relapsing-remittent (RR), primary-progressive (PP), and secondary-progressive (SP). Few patients (85%) were earlier diagnosed as CIS during the clinical exam. CIS patients presents then RR. RR can shift later to SP with or without superimposed relapses [Lawton et al., 2015]. 15% of MS patients presents a PP, that is characterized by a continuous worsening of symptoms without relapses since diagnosis [Lublin, 2014].

In addition to clinical examination, Magnetic Resonance Imaging (MRI) helps the diagnosis and the monitoring of disease evolution. Conventional MRI (such as T1-weighted and T2-weighted imaging) has important specificity and sensitivity in detecting pathological tissue damage and obtaining valuable predictive information on disease evolution [Filippi, 2001; Fisniku et al., 2008]. In this study, we used the lesion load and the grey matter volume because they reflect better the levels of neuroaxonal loss and demyelination [Peterson et al., 2001; Minneboo et al., 2009]. However, advanced MRI (such as Diffusion Tensor Imaging (DTI) and Spectroscopic Imaging) has a higher specificity in detecting microstructural damages in white matter than conventional imaging [Ge et al., 2004; Sbardella et al., 2013] because they provide complementary information giving the diffusion alteration and metabolic change [Filippi et al., 2001; Filippi, 2001; Filippi, 2001]. In addition, as the

patients in different subtypes show different diffusivity patterns, DTI is able to separate the MS subtypes [Sbardella et al., 2013]. Therefore, we propose to use DTI data jointly with the conventional imaging. The fractional anisotropy (FA) measure was chosen among DTI measurements because it is more sensitive in detecting microscopic changes related to inflammation [Hannoun et al., 2012].

MS affects especially the white matter of the brain and the lesions appear mostly around the ventricles and periventricular white matter. The floor of the lateral ventricles forms a region called Corpus Callosum (CC). So, CC is the one of the regions the most frequently affected by lesions in MS [Barnard, 1974; Gean-Marton et al., 1991; Ge et al., 2006]. In 93% of MS patients, the lesions are found in CC [Gean-Marton et al., 1991]. Moreover, CC is the major myelinated bundle of the brain providing the connection between cortical and subcortical regions of the two brain hemispheres. For this reason, the impact of lesions in CC is more severe in comparison with the lesions in lobar white matter.

Furthermore, several studies have demonstrated that callosal changes, measured with DTI, were correlated with cognitive and physical disability [Sigal et al., 2012, Yaldizli et al., 2010; Rimkus et al., 2010; Llufríu et al., 2012]. FA measurements were significantly lower in rostrum, body, and splenium part of CC compared to control group [Hasan et al., 2005; Rueda et al., 2008; Warlop et al., 2008]. The association between EDSS scores and CC atrophy measured with conventional MR was approved in some studies [Hasan et al., 2005; Rueda et al., 2008; Warlop et al., 2008; Schreiber et al., 2001]. However, no significant correlation was found between disability and callosal atrophy measured with conventional imaging in RR patients [Barkhof et al., 1998]. This contrast can be the result of insufficiency of conventional imaging on CC and the use of advanced MRI techniques -as DTI- can be more suitable for CC data in MS patients. For this reason, the FA measured in CC was used in this study.

Among the studies interested in MS progression, many articles have used logistic regression to predict the presence or absence of progression (at least 1 point increase in EDSS score) or ordinal logistic regression to predict the EDSS step changes in ordinal categories [Sastre-Garriga et al., 2005; Minneboo et al., 2008, Khaleeli et al., 2008]. The predictive ability of the clinical and imaging data for the disability evolution in MS is analyzed with linear regression models [Minneboo et al., 2009; Furby et al., 2010; Bodini et al., 2011; Enzinger et al., 2011; Popescu et al., 2013]. At the same time, the multilevel approach is performed to consider individual alteration during disability progression [Di Serio et al. 2009; Lawton et al., 2015]. This latter model was able to include heterogeneity among individuals. However, the evolution of disability in MS does not show a unique mean evolution profile, thus, it would be interesting to consider heterogeneity originated from different mean evolution profiles of the MS patients.

The principal aim of this part of the thesis is to develop a generalizable predictive model of disability evolution of MS patients considering unobserved subgroups (different mean-evolution profiles). For this aim, the latent class linear mixed model was used to predict EDSS evolution over 5 years on the basis of the clinical, biological and imaging data taken at study onset.

MATERIALS

Clinical data

Eighty patients fulfilling the Mac Donald criteria were included in a standardized clinical and MRI protocol within the frame of the AMSEP project in Lyon Neurological Hospital. This population was divided into four groups of MS patients depending on the clinical form: CIS (n=12), RR (n=27), SP (n=16), and PP (n=25). The patients were followed up with standardized clinical and MRI examination every six months during the first three years then at one year intervals during two years. Clinical examination provided the Expanded Disability Status Scale (EDSS), Multiple Sclerosis Functional Composite (MSFC) with its three dimensions: Timed 25 Foot Walk (T25FW), 9-Hole Peg Test (9HPT) and Paced Auditory Serial Addition Test (PASAT), disease duration at study onset. Age of the patients was also provided. Five patients were excluded from the AMSEP cohort because one or more MSFC component(s) could not be measured at study onset.

Image Acquisition and Processing

MS patients underwent an MR examination on a 1.5T Siemens Sonata system (Siemens Medical Solution, Erlangen, Germany) using an 8-channel head-coil. The MR protocol consisted in the acquisition of a sagittal 3D-T1 sequence ($1 \times 1 \times 1 \text{ mm}^3$, TE/TR = 4/2 000 ms) and an axial 2D-spin-echo DTI sequence (TE/TR = 86/6900 ms; 2×24 directions of gradient diffusion; $b = 1000 \text{ s.mm}^{-2}$, spatial resolution of $2.5 \times 2.5 \times 2.5 \text{ mm}^3$ oriented in the AC-PC plane).

The lesion load was measured with FLAIR sequence and white and gray matter volumes were measured with T1 imaging without gadolinium contrast agent.

METHODS

The latent class linear mixed model

We used the latent class linear mixed model, an extension of the linear mixed model used for longitudinal outcomes. Let $Y_i = (Y_{i1}, \dots, Y_{in_i})$ be the n_i repeated outcome measures of subject i in a sample of N subjects. The linear mixed model [Laird et al., 1982] is defined as:

$$Y_i = X_i\beta + Z_iu_i + \epsilon_i$$

where X_i is a $n_i \times p$ matrix is related to the vector of fixed effects β in length p , Z_i is a $n_i \times q$ design matrix is related to the vector of random effects u_i in length q . The linear mixed models assume that the population is homogeneous and that the random effects are normally distributed as $u_i \sim N(0, \mathbf{B})$. ϵ_i are the measurement errors that are assumed to be normally distributed with mean zero and covariance matrix $\sigma^2 I_{n_i}$ and to be independent from the random effects u_i .

However, the latent class linear mixed model considers that the population is not homogeneous and consists of G subgroups also called latent classes. Each latent class shows different distributions with the class-specific matrix of variance-covariance. The latent class mixed model is defined as:

$$Y_i = X_{1i}\beta + X_{2i}\nu_g + Z_iu_{ig} + \epsilon_i \quad (1)$$

where X_{1i} and X_{2i} are the covariate matrices and associated with the vector of fixed effects, at population and latent class levels, respectively. Z_i is the class-specific covariate matrix of class g and is associated with the random effects $u_{ig} \sim N(\mu_g, \mathbf{B}_g)$ where μ_g and \mathbf{B}_g are, respectively, the class-specific mean and the covariance matrix.

Each subject can belong to one latent class c_i and the probability to be in class c_i is defined using a multinomial logistic model as:

$$\pi_{ig} = P(c_i = g | X_{ci}) = \frac{e^{\xi_{0g} + X_{ci}\xi_{1g}}}{\sum_{l=1}^G e^{\xi_{0l} + X_{ci}\xi_{1l}}} \quad (2)$$

where ξ_{0g} is the intercept of the latent class g , ξ_{1g} is the class specific parameter and X_{ci} represents the individual class-specific covariate matrix. Moreover, the sum of probabilities of being in various classes is equal to one ($\sum_{g=1}^G \pi_{ig} = 1$ for $0 \leq \pi_{ig} \leq 1$).

Parameter estimation

Parameter estimation in the latent class linear mixed model can be performed with the maximum likelihood. Let θ_G be the vector of parameters of the model and L_i the individual contribution to the likelihood of the model. The individual contribution to the likelihood of the latent class linear mixed model is defined as:

$$L_i(\theta_G) = \sum_{g=1}^G \pi_{ig} \times \phi_{ig}(Y_i | c_i = g; \theta_G) \quad (3)$$

where π_{ig} is the probability of belonging to class g and ϕ_{ig} is the density function of a multivariate normal distribution with mean $X_{1i}\beta + X_{2i}\nu_g$ and variance $Z_i B Z_i^T + \sigma^2 I_{n_i}$.

The estimates of θ_G can be obtained with vector $\hat{\theta}_G$ that maximizes the log-likelihood defined as $l(\theta_G) = \sum_{i=1}^N \log(L_i(\theta_G))$. Parameter estimation by iterative algorithm is possible with the EM family [Verbeke et Lesaffre, 1996], the Newton-Raphson family [Proust et Gadda, 2005], or the Marquardt algorithm [Marquardt, 1963]. Here, we used the Marquardt algorithm for parameter estimation because of its speed and convergence rate [Proust-Lima et al., 2015].

Posterior classification

After parameter estimation, the posterior probability of assigning each subject to a latent class g can be obtained using following equation:

$$\widehat{\pi}_{ig} = P(c_i = g | X_{1i}, X_{ci}, Y_i, \hat{\theta}_G) = \frac{\pi_{ig} \times \phi_{ig}(Y_i | c_i = g; \theta_G)}{\sum_{l=1}^G \pi_{il} \times \phi_{il}(Y_i | c_i = l; \theta_G)} \quad (4)$$

The subject is assigned to the latent class that maximizes the posterior probability defined as: $\hat{c}_i = \text{argmax}_g(\widehat{\pi}_{ig})$. This assignment allows obtaining the posterior classification that can then refer to the goodness-of-fit of the model [Proust-Lima et al. 2015]. The higher is the mean of the posterior probabilities obtained at each latent class, the better is the classification.

Longitudinal predictions

When the number of latent classes is higher than one, the class-specific subject-specific predictions for individual i at j a given time point can be calculated as:

$$\widehat{Y}_{ijg}^{(SS)} = X_{1i} \hat{\beta} + X_{2i} \hat{v}_g + Z_i \hat{u}_{ig}$$

where $\hat{\beta}$ and \hat{v}_g are the estimated parameters of respectively the fixed and the class-specific fixed effects. \hat{u}_{ig} are the empirical Bayes estimates for random effects.

Application of latent class linear mixed model

All statistical analyses were performed with *lcmm* package in R software version 3.4.0 (2017-04-21).

Function *hlme* in **lcmm** package of software R was used to implement the latent class linear mixed model. There are three main arguments of the *hlme* function:

1. **fixed** argument: contains two-sided formula including the outcome (dependent variable) and the independent variable(s) which have a common effect on all individuals overall latent classes.
2. **mixture** argument: indicates the variable(s) that have a specific effect on each latent class.
3. **random** argument: contains the variable(s) that have a specific effect at the individual level.

The EDSS score was used as the outcome in the models. The independent variables were clinical and imaging data such as the time, age, disease duration, T25FW, 9HPT, and grey matter volume, lesion load, and fractional anisotropy. The EDSS score and time were used as longitudinal data. The other clinical and imaging variables were entered into the model with their values at study onset.

Before fitting the model, the latent class mixed model requires setting the number of latent class, the variable that determines the latent classes (in the mixture argument) and the time function (linear, quadratic, and square root). Time, age, disease duration, T25FW and 9HPT, grey matter volume, lesion load, and fractional anisotropy were tested separately in the mixture argument with the linear, quadratic, and square root of time in three and four latent classes. So, we established 48 models and these models were compared using the Bayesian information criterion (BIC). The model with the

lower BIC value provided the best combination of number of latent class, variable that determines the latent classes and time function.

All clinical and imaging variables were used in the fixed argument during the choice of the best combination. Besides, age at study onset was used as the variable that exerts the individual effect (in the random argument) in all models because there was a great inter-patient variability in terms of age at study onset.

First, the clinical model was established with age, disease duration, T25FW, and 9HPT at study onset, and time in the fixed argument. Second, the imaging model was established with GMV, LL, and FA in the fixed argument. Third, the imaging variables were added one by one or two by two or all three together with the clinical variables into the clinical model to obtain the combined models. Finally, we compared the three clinical, imaging and combined models. Model comparison was performed mainly using the BIC criterion. A low BIC indicates a better fit of the model, however, the mean of the posterior probabilities was also reported in this study to examine the goodness of fit and allow result comparisons with previous studies.

Statistical analyses

Because most data did not follow the normal distribution, medians, 1st, and 3rd quartiles were used to describe the data. Consequently, non-parametric statistical tests were used to examine differences in clinical and imaging data among the patients of the four clinical subtypes. The Kruskal-Wallis test was used first to analyze whether the distributions of subtypes were significantly different. In case of significant difference, a Mann-Whitney U Test was used to compare subtype medians two by two. Statistical significance was considered at $p < 0.05$.

RESULTS

The model with three latent classes with time as the variable with a specific effect on each latent class (in the mixture argument), and with a linear time function (in the fixed argument) gave the lowest BIC value (See Appendix). Therefore, the models that include time in the mixture argument and age in the random argument were built for the following sections.

Clinical model

In the clinical model, all clinical variables had significant effects on the evolution of the EDSS score ($\beta_{disease\ duration}=0.11$, $\beta_{age}=0.08$, $\beta_{T25FW}=0.12$, and $\beta_{9HPT}=0.07$ where p-value < 0.05 for all variables). The effect of time was presented for each latent class and its effect was significant in two classes ($\beta_{time,class\ 1}=0.05$ and $\beta_{time,class\ 3}=-0.09$ where p-value < 0.05). These classes were the latent classes (class 1 and 3) which contain fewer patients than the other latent classes (class 2) (12% and 4% vs 84%). However, time had no significant effect in class 2 (p-value=0.06).

Variable	Coef	Se	Wald	p-value
Disease Duration	0.10772	0.02392	4.504	0.00001
Age	0.08036	0.01907	4.215	0.00003
T25M	0.12060	0.03990	3.022	0.00251
9HPT	0.06453	0.01708	3.779	0.00016
Time class1	0.04715	0.00436	10.805	0.00000
Time class2	0.00274	0.00148	1.858	0.06321
Time class3	-0.09571	0.01254	-7.631	0.00000

Table 1 - The variables used in the fixed and mixture arguments of the clinical model. Coef: Parameter coefficient, Se: Standard Error, and Wald: Statistical test value.

Figure 1 shows the predicted points and observed mean trajectories obtained with the clinical model. The graph shows that there is one class with stable progression (Class 2). However, Classes 1 and 3 showed respectively severe and moderate evolutions. The predicted points were close to the observed mean trajectories for classes 1 and 2. Classes 1 and 2 included respectively 9 and 63 patients. However, Class 3 contains 3 patients. The fewer number of patients in Class 3 might cause the greater difference between the predicted and observed mean trajectories.

Table 2 shows the classification results obtained with the clinical model. Most of the patients were found in latent class 2 (84%) which had a stable mean evolution. Latent class 1 included RR, PP, and SP patients with more severe clinical evolution than CIS. Contrary to latent class 1, latent class 3 included patients with moderate evolution and there were two CIS patients and one RR patient.

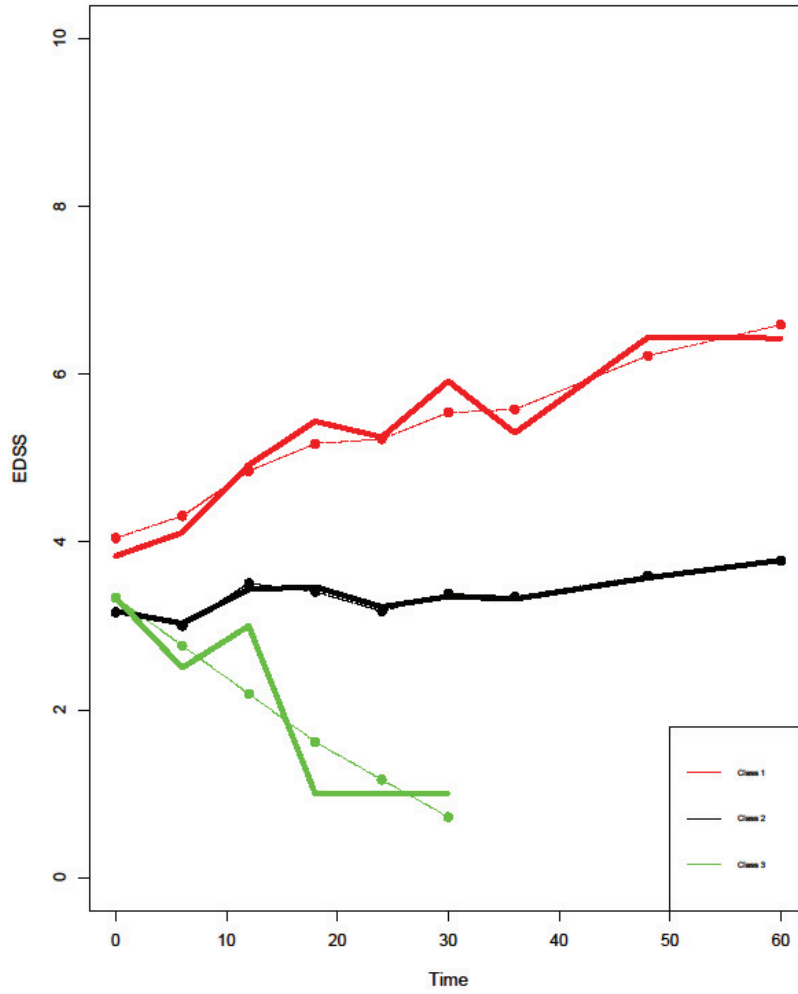


Figure 1. The solid lines show the observed mean trajectories and the dotted lines their 95% confidence boundaries. The points show the mean predicted evolution of each class obtained with clinical model.

	Class 1	Class 2	Class 3
CIS	0	10	2
RR	2	23	1
PP	2	13	0
SP	5	17	0

Table 2 - Classification obtained with the clinical model and all MS subtypes.

The patients in the three latent classes were compared in terms of age, disease duration, T25FW, 9HPT, GMV, LL, and FA. The number of patients, the median, 1st, and 3rd quartiles are given in Table 3. There was a significant difference between the three latent classes in terms of disease duration at study onset (p-value=0.023). The medians of disease duration at study onset of latent classes 1 and 2 were equal to 6.6 and 5.9 months, respectively, whereas the median of disease duration of latent class 3 was equal to 0.3 months. The other clinical and imaging variables were not significantly different among the three latent classes (p-value>0.05).

Variable	Class 1	Class 2	Class 3	p-value
	N=9	N=63	N=3	
Age	40.5 (32.4, 44.0)	39.1 (30.5, 42.8)	34.4 (28.5, 38.7)	0.71
Disease Duration	6.6 (4.9, 10.5)	5.9 (3.4, 9.0)	0.3 (0.2, 0.8)	0.02
9HPT	24.8 (20.1, 32.6)	22.0 (19.3, 26.1)	23.4 (21.2, 26.3)	0.70
25FW	5.1 (4.1, 7.0)	5.4 (4.5, 6.6)	4.2 (3.8, 4.3)	0.08
GMV	831.1 (799.8, 850.7)	824.7 (800.7, 858.0)	869.8 (843.1, 875.9)	0.53
LL	26.5 (15.1, 35.4)	15.1 (6.1, 32.9)	7.2 (7.1, 7.4)	0.15
FA	0.12 (-0.20, 0.73)	0.41 (0.11, 0.78)	0.43 (0.41, 0.74)	0.57

Table 3 - Clinical and imaging variables according to each latent class obtained with the clinical model. Median (1st, 3rd quartile). The p-value is the result of Kruskal-Wallis test which was performed to compare the medians of variables.

Imaging model

Table 4 shows the result of the imaging model established with GMV, LL, and FA in the fixed argument. The effect of the GMV was significantly different than 0 (p-value<0.05). A decrease of GMV value causes an increase in the clinical disability ($\beta_{GMV} = -0.02$). However, LL and FA did not show significant effects on disease evolution measured with EDSS score in the imaging model (p-value=0.26).

Variable	Coef	Se	Wald	p-value
GMV	-0.01625	0.00404	-4.024	0.00006
LL	-0.00049	0.01243	-0.039	0.96880
FA	-0.26589	0.23815	-1.116	0.26421
9HPT	0.06453	0.01708	3.779	0.00016
Time class1	0.00522	0.00167	3.130	0.00175
Time class2	0.04732	0.00391	12.108	0.00000
Time class3	-0.02096	0.00520	-4.027	0.00006

Table 4 - The variables used in fixed and mixture arguments in the combined model. Coef: Parameter coefficient, Se: Standard Error, and Wald: Statistical test value.

Figure 2 shows the predictions and observed mean trajectories obtained with the imaging model. The predicted trajectories were similar to the trajectories obtained with the clinical model. Class 2 trajectory included the majority of patients (73%) and showed a stable evolution. Class 1 included two RR, two PP, and four SP patients. The predicted mean trajectory of these patients showed severe evolutions. However, Class 3 showed a moderate evolution with decrease over time. This cluster included the patients who have a benign evolution (CIS (n=5) and RR (n=7)).

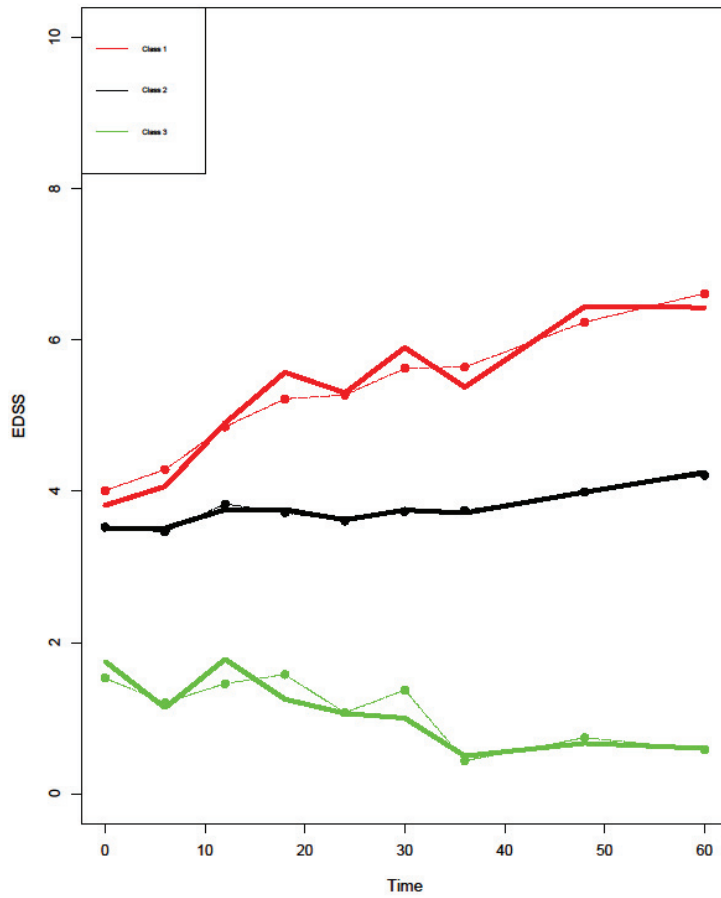


Figure 2. The solid lines show the observed mean trajectories and the dotted lines their 95% confidence boundaries. The points show the mean predicted evolution of each class obtained with imaging model.

	Class 1	Class 2	Class 3
CIS	0	7	5
RR	2	17	7
PP	2	13	0
SP	4	18	0

Table 5. Classification obtained with the imaging model with all MS subtypes.

Combined model with clinical and imaging variables

Table 6 shows the BIC, the mean of posterior probabilities, the log-likelihood, and the AIC of the clinical, the imaging, and the combined model built with seven different combinations of imaging variables. The highest BIC was obtained with the imaging model and the lowest BIC with the clinical model. All models had mean posterior probabilities over 0.89. Most of the combined models had a greater mean posterior probability than the clinical and imaging model. The log-likelihood ranged between -542 and -505. The model with clinical, GMV and FA variables and the model with the clinical and all imaging variables showed the best likelihood values. AIC values ranged between 1047 and 1115. The model with clinical, GMV and FA variables had the lowest AIC value.

Model	BIC	Posterior Probability	Loglik	AIC
Clinical	1086	0.970	-508	1049
Imaging	1149	0.949	-542	1115
Clinical +GMV	1105	0.990	-516	1066
Clinical +LL	1111	0.890	-519	1072
Clinical +FA	1088	0.985	-507	1048
Clinical +GMV+LL	1123	0.982	-522	1081
Clinical +GMV+FA	1089	0.989	-505	1047
Clinical + LL+FA	1092	0.985	-507	1050
Clinical +GMV+LL+FA	1093	0.989	-505	1049

Table 6 - BIC, mean of posterior probabilities, Log-likelihood, and AIC results with the clinical, the imaging and the combined models built with different combinations of imaging variables.

Table 7 shows the results of the combined model built with clinical and all three imaging parameters. The effects of the clinical variables were still significant ($\beta_{\text{disease duration}}=0.11$, $\beta_{\text{age}}=0.06$, $\beta_{\text{T2FW}}=0.13$, and $\beta_{\text{9HPT}}=0.04$ where $p\text{-value}<0.05$ for all). However, the effects of the imaging variables were not significantly different from zero ($p\text{-value}>0.05$).

Variable	Coef	Se	Wald	p-value
Disease duration	0.10505	0.02361	4.450	0.00001
Age	0.05980	0.02244	2.664	0.00771
T25FW	0.12681	0.03747	3.384	0.00071
9HPT	0.04244	0.01900	2.233	0.02552
GMV	-0.00751	0.00407	-1.845	0.06502
LL	-0.00372	0.01187	-0.314	0.75381
FA	-0.25584	0.21576	-1.186	0.23571
Time class1	-0.09566	0.01256	-7.616	0.00000
Time class2	0.00285	0.00140	2.045	0.04090
Time class3	0.04754	0.00380	12.522	0.00000

Table 7 - The variables used in fixed and mixture arguments in the combined model. Coef: Parameter coefficient, Se: Standard Error, and Wald: Statistical test value.

Figure 3 shows predicted and observed mean trajectories obtained with the combined model built with GMV, LL, and FA together. The observed mean trajectories and predicted points were almost similar to the result obtained with the clinical data. Moreover, the classification performed with the combined model (see Table 8) did not significantly change the classification obtained with the clinical model. There was one more patient classified in latent Class 2 (with stable evolution) instead of latent Class 1 (with severe evolution).

As only one patient changed class with the combined model, the comparison results of the combined model among latent classes did not change compared to the comparison results of the clinical model. This means that the disease duration at study onset was significantly different among three latent classes (p -value=0.028). However, the three latent classes were not different considering other clinical and imaging variables (p -value>0.05).

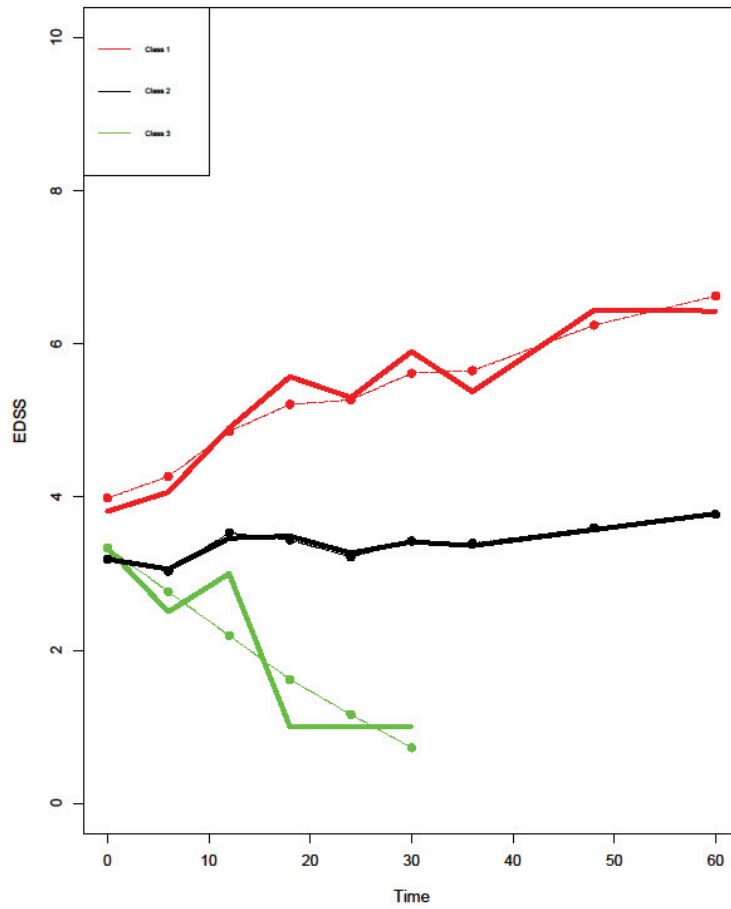


Figure 3. The solid lines show the observed mean trajectories and the dotted lines their 95% confidence boundaries. The points show the mean predicted evolution of each class performed with combined model.

	Class 1	Class 2	Class 3
CIS	0	10	2
RR	2	23	1
PP	2	13	0
SP	4	18	0

Table 8. Classification obtained with the combined model

(clinical variables with GMV, LL and FA) for all MS subtypes.

DISCUSSION

The latent class linear mixed model was used to model the evolution of disability in patients with multiple sclerosis, thus, to predict individual disability evolution measured with EDSS score. To our knowledge, this is the first study that predicts long-term disability evolution considering unobserved subgroups of multiple sclerosis patients of all clinical subtypes. First, the clinical and imaging models were separately built with, respectively, clinical and imaging variables. Then, multimodal imaging variables were added to the clinical model to obtain combined models. As per the BIC criterion, the clinical model had a higher predictive accuracy in comparison with the imaging or the combined models considering the BIC criterion.

In the clinical and combined models, all clinical variables had a significant predictive effect on disease evolution. Previous studies showed also age and disease duration were significant in predictive modeling [Confavreux et al., 2003; Scalfari et al., 2011]. In addition, 9HPT and T25FW were considered as two of the best measures across a wide range of indicators of MS disability [Kieseier & Pozzilli, 2011; Kraft et al., 2014]. However, no study was yet performed to check whether the 9HPT had a significant effect on disability evolution. On the other hand, an early change in T25FW had a predictive value on the long-term EDSS evolution in progressive patients (PP and SP) [Bosma et al., 2012]. Our results confirm that the effects of T25FW and 9HPT are significant in all clinical subtypes.

Previous studies on predictive modeling of disability evolution in patients with multiple sclerosis used logistic regression models, Kaplan-Meier analyses, Markov models, multilevel modeling, and, mostly, linear mixed models [Popescu et al., 2013; Minneboo et al., 2009; Bodini et al., 2011; Minneboo et al., 2007; Palace et al., 2014; Lawton et al., 2015; Confavreux et al., 2003; Confavreux et al., 2000]. A recent study used latent class linear mixed models in PP patients to identify unobserved classes [Signori et al., 2017], selected the best model with the BIC criterion, and identified three subgroups

of PP patients. However, our study remains the first one to use latent class mixed models, consider unobserved classes, and include all clinical subtypes of MS.

In our study, the multiple sclerosis patients were classified into three latent classes having a stable, a severe and a moderate evolution over time. Most of the patients (84%) showed no progression of disability over five years after the study onset. As MS is a long-lasting disease, the five years might be insufficient to show disability worsening. However, 12% of the patients were assigned to the class with severe disability progression and 4% to the class with moderate evolution. Assigning patients to either of these two latent classes may help treatment decision. For example, patients with a probable severe evolution would benefit from second-line therapies.

EDSS score 4 is known as the threshold of limited walking disability though a patient may be able to walk more than 500m. By modeling, the predicted EDSS score values at the study onset were around EDSS score 4 in the three latent classes then each class evolved differently. Thus, our predictions were able to identify the patients who might have a severe, stable or moderate evolution after reaching the threshold of limited walking disability.

In this study, GMV had a significant effect in the imaging model. Previous studies have shown that the effect of gray matter volume was significant in increasing the risk of severe disability of MS patients [Rovaris et al., 2006; Fisniku et al., 2008, Rosendaal et al., 2011; Sepulcre et al., 2006]. However, the effects of the multimodal imaging variables were not statistically significant in the combined models. Enzinger et al. have performed a multivariate analysis and found that no MRI variables had a significant effect on disability evolution [Enzinger et al., 2011]. A study performed on the predictive value of the lesion load showed that the lesion load observed in the gray and white matter was a strong predictor of the 10-year EDSS in MS [Popescu et al., 2013]. However, here, the lesion load was found significant neither in the imaging and nor in the combined models.

One limitation of the present work was the use of the lesion load measured in the whole brain and, not locally. The location and size of the lesion are very important markers of the clinical disability. A future modeling work on the evolution of clinical disability may consider the focal lesion load and, gray and white matter loss.

To conclude on this issue, the latent class linear mixed model allowed building a well-fitted predictive model for disability evolution in patients with MS and this model showed high accuracy results. The model developed here is highly promising in predicting individual long-term disability evolution in all clinical subtypes of MS.

REFERENCES

1. Goldenberg MM. Multiple sclerosis review. *P T*. 2012;37(3):175-184.
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3351877&tool=pmcentrez&rendertype=abstract>.
2. Lawton M, Tilling K, Robertson N, et al. A longitudinal model for disease progression was developed and applied to multiple sclerosis. *J Clin Epidemiol*. 2015;68(11):1355-1365.
doi:10.1016/j.jclinepi.2015.05.003.
3. Lublin FD. New multiple sclerosis phenotypic classification. *Eur Neurol*. 2014.
doi:10.1159/000367614.
4. Filippi M. Non-conventional MR techniques to monitor the evolution of multiple sclerosis. 2001:195-200.
5. Fisniku LK, Brex PA, Altmann DR, et al. Disability and T 2 MRI lesions : a 20-year follow-up of patients with relapse onset of multiple sclerosis. 2008:808-817. doi:10.1093/brain/awm329.
6. Minneboo A, Uitdehaag BMJ, Jongen P, Vrenken H, Knol DL, Walderveen MAA Van. Association between MRI parameters and the MS severity scale : a 12 year follow-up study. 2009;(December 2008):632-637.
7. Ge Y, Law M, Johnson G, et al. Preferential Occult Injury of Corpus Callosum in Imaging. 2004;7:1-7. doi:10.1002/jmri.20083.
8. Sbardella E, Tona F, Petsas N, Pantano P. DTI Measurements in Multiple Sclerosis : Evaluation of Brain Damage and Clinical Implications. *Mult Scler Int*. 2013;2013:1-11.
9. Filippi M, Cercignani M, Inglese M, Horsfield MA, Comi G. Diffusion tensor magnetic resonance. 2001.
10. Filippi M. Magnetic resonance imaging findings predicting subsequent disease course in patients at presentation with clinically isolated syndromes suggestive of multiple sclerosis. *Neurol Sci*. 2001;22(SUPPL. 2):51-53. doi:10.1007/s100720100033.

11. Hannoun S, Durand-dubief F, Confavreux C, et al. Diffusion Tensor – MRI Evidence for Extra-Axonal Neuronal Degeneration in Caudate and Thalamic Nuclei of Patients with Multiple Sclerosis. *AJNR*. 2012;33:1363-1368.
12. Ge Y. REVIEW ARTICLE Multiple Sclerosis : The Role of MR Imaging. 2006:1165-1176.
13. Barnard RO, Triggs M. Corpus callosum in multiple sclerosis. *J Neurol*. 1974;37:1259-1264.
14. Gean-Marton A, Vezina G, Marton KI, et al. Abnormal Corpus Callosum : A sensitive and Specific Indicator of Multiple Sclerosis.
15. Rimkus CDM, Junqueira TDF, Lyra KP, et al. Corpus Callosum Microstructural Changes Correlate with Cognitive Dysfunction in Early Stages of Relapsing-Remitting Multiple Sclerosis : Axial and Radial Diffusivities Approach Maria Concepci on. 2011;2011(Cc). doi:10.1155/2011/304875.
16. Llufríu S, Blanco Y, Martínez-heras E, et al. Influence of Corpus Callosum Damage on Cognition and Physical Disability in Multiple Sclerosis : A Multimodal Study. 2012;7(5):1-7. doi:10.1371/journal.pone.0037167.
17. Sigal T, Shmuel M, Mark D, Gil H, Anat A. Diffusion Tensor Imaging of Corpus Callosum Integrity in Multiple Sclerosis : Correlation with Disease Variables. 2012;22:33-37. doi:10.1111/j.1552-6569.2010.00556.x.
18. Yaldizli O, Atefy R, Gass A, et al. Corpus callosum index and long-term disability in multiple sclerosis patients. *J Neurol*. 2010;257:1256-1264. doi:10.1007/s00415-010-5503-x.
19. Hasan KM, Gupta RK, Santos RM, Wolinsky JS, Narayana PA. Diffusion Tensor Fractional Anisotropy of the Normal-Appearing Seven Segments of the Corpus Callosum in Healthy Adults and Relapsing- Remitting Multiple Sclerosis Patients. 2005;743:735-743. doi:10.1002/jmri.20296.
20. Rueda F, Celso L, Jr H, Papais-alvarenga RM, Gasparetto EL. Diffusion tensor MR imaging evaluation of the corpus callosum of patients with multiple sclerosis. 2008;66(May):449-453.
21. Warlop NP, Achten E, Debruyne J, Vingerhoets G. *Neuropsychologia* Diffusion weighted

- callosal integrity reflects interhemispheric communication efficiency in multiple sclerosis. 2008;46:2258-2264. doi:10.1016/j.neuropsychologia.2008.02.010.
22. Schreiber K, Ps S, Wagner A. Correlations of brain MRI parameters to disability in multiple sclerosis. 2001:24-30.
 23. Barkhof F, Filippi M, Miller DH, et al. Comparison of MRI criteria at first presentation to predict conversion to clinically definite multiple sclerosis. *Brain*. 1997:2059-2069.
 24. Sastre-garriga J, Ingle GT, Jasperse B, et al. Long-term clinical outcome of primary progressive MS : Predictive value of clinical and MRI data. *Neurology*. 2005:633-635.
 25. Minneboo A, Jasperse B, Barkhof F, et al. Predicting short-term disability progression in early multiple sclerosis : added value of MRI parameters. *JNNP*. 2008;79:917-924. doi:10.1136/jnnp.2007.124123.
 26. Khaleeli Z, Ciccarelli O, Manfredonia F, et al. Predicting Progression in Primary Progressive Multiple Sclerosis : A 10-Year Multicenter Study. 2008:790-793. doi:10.1002/ana.21375.
 27. Furby J, Hayton T, Altmann D, et al. A longitudinal study of MRI-detected atrophy in secondary progressive multiple sclerosis. *J Neurol*. 2010. doi:10.1007/s00415-010-5563-y.
 28. Bodini B, Battaglini M, De Stefano N, et al. T2 lesion location really matters: a 10 year follow-up study in primary progressive multiple sclerosis. doi:10.1136/jnnp.2009.201574.
 29. Enzinger C, Fuchs S, Pichler A, et al. Predicting the severity of relapsing-remitting MS: The contribution of cross-sectional and short-term follow-up MRI data. doi:10.1177/1352458510394454.
 30. Popescu V, Agosta F, Hulst HE, et al. Brain atrophy and lesion load predict long term disability in multiple sclerosis. *J Neurol Neurosurg Psychiatry*. 2013;84:1082-1091. <http://dx>.
 31. Serio C Di, Lamina C. Investigating Determinants of Multiple Sclerosis in Longitudinal Studies : A Bayesian Approach. 2009;2009. doi:10.1155/2009/198320.
 32. Proust C, Jacqmin-Gadda H. Estimation of linear mixed models with a mixture of distribution for the random effects. *Comput Methods Programs Biomed*. 2005;78:165-173.

doi:10.1016/j.cmpb.2004.12.004.

33. Verbeke G, Lesaffre E. A Linear Mixed-Effects Model With Heterogeneity in the Random-Effects Population. 1996;91(433). doi:10.1080/01621459.1996.10476679.
34. Marquardt D. An algorithm for least-squares estimation of nonlinear parameters. *J Soc Indust Appl Math*. 1963;11(2):431-441.
35. Proust-lima C, Philipps, VivianeLiquet B. Estimation of Extended Mixed Models Using Latent Classes and Latent Processes : The R Package lcmm. 2015.
36. Scalfari A, Neuhaus A, Daumer M, Ebers GC, Muraro PA. Age and disability accumulation in multiple sclerosis. *Neurology*. 2011;77:1246–1252.
37. Confavreux C, Vukusic S, Adeleine P. Early clinical predictors and progression of irreversible disability in multiple sclerosis: An amnesic process. *Brain*. 2003. doi:10.1093/brain/awg081.
38. Kieseier BC, Pozzilli C. Assessing walking disability in multiple sclerosis. *Mult Scler J*. 2012;18(7):914-924. doi:10.1177/1352458512444498.
39. Kraft GH, Amtmann D, Bennett SE, et al. Assessment of Upper Extremity Function in Multiple Sclerosis : Review and Opinion. *Postgrad Med*. 2014;126(5). doi:10.3810/pgm.2014.09.2803.
40. Bosma LVAE, Kragt JJ, Knol DL, Polman CH, Uitdehaag BMJ. Clinical scales in progressive MS : predicting long-term disability. *Mult Scler J*. 2012;18(3):345-350. doi:10.1177/1352458511419880.
41. Rovaris M, Judica E, Gallo A, et al. Grey matter damage predicts the evolution of primary progressive multiple sclerosis at 5 years. *Brain*. 2006;129:2628-2634. doi:10.1093/brain/awl222.
42. Roosendaal SD, Bendfeldt K, Vrenken H, et al. Grey matter volume in a large cohort of MS patients : relation to MRI parameters and disability. *Mult Scler J*. 2011;17(9):1098-1106. doi:10.1177/1352458511404916.
43. Sepulcre J, Sastre-garriga J et al. Regional Gray Matter Atrophy in Early Primary Progressive Multiple Sclerosis, A Voxel-Based Morphometry Study. *ARCH NEUROL*. 2006;63.

44. Popescu V, Agosta F, Hulst HE, et al. Brain atrophy and lesion load predict long term disability in multiple sclerosis. 2013;1082-1091. doi:10.1136/jnnp-2012-304094.
45. Bodini B, Battaglini M, Stefano N De, et al. T2 lesion location really matters : a 10 year follow-up study in primary progressive multiple sclerosis. *J Neurol Neurosurg Psychiatry*. 2011;82:72-77. doi:10.1136/jnnp.2009.201574.
46. Minneboo A, Jasperse B, Barkhof F, et al. Predicting short-term disability progression in early multiple sclerosis: added value of MRI parameters. 2007. doi:10.1136/jnnp.2007.124123.
47. Palace J, Bregenzer T, Tremlett H, et al. UK multiple sclerosis risk-sharing scheme : a new natural history dataset and an improved Markov model. 2014. doi:10.1136/bmjopen-2013-004073.
48. Confavreux C, Vukusic S, Adeleine P, Lyon HC De. Early clinical predictors and progression of irreversible disability in multiple sclerosis : an amnesic process. 2003. doi:10.1093/brain/awg081.
49. Confavreux C, Vukusic S, Moreau T, Adeleine P. Relapses and progression of disability in multiple sclerosis. *N Engl J Med*. 2000;343(20):1430-1438.

APPENDIX

Supplementary Table 1 shows the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) results obtained with 48 models fitted with 8 different variables (time, age, disease duration, 9HPT, and T25FW) in mixture argument, and 3 types of time function (linear, square root and polynomial) in 3 and 4 latent classes.

Variable		Time function	BIC	AIC
Time	3	Linear	1093	1049
		Square Root	1097	1051
		Polynomial	1121	1074
	4	Linear	1100	1049
		Square Root	1099	1052
		Polynomial	1110	1056
Age	3	Linear	1207	1163
		Square Root	1212	1166
		Polynomial	1212	1166
	4	Linear	1220	1169
		Square Root	1224	1171
		Polynomial	1224	1170
Disease Duration	3	Linear	1221	1176
		Square Root	1207	1161
		Polynomial	1214	1167
	4	Linear	1217	1166
		Square Root	1224	1171
		Polynomial	1220	1167
9HPT	3	Linear	1215	1171

		Square Root	1223	1177
		Polynomial	1209	1163
	4	Linear	1222	1171
		Square Root	1224	1177
		Polynomial	1216	1169
T25FW	3	Linear	1207	1162
		Square Root	1212	1166
		Polynomial	1210	1164
	4	Linear	1200	1149
		Square Root	1202	1149
		Polynomial	1218	1164
GMV	3	Linear	1209	1165
		Square Root	1210	1164
		Polynomial	1206	1160
	4	Linear	1219	1168
		Square Root	1212	1159
		Polynomial	1216	1163
LL	3	Linear	1207	1163
		Square Root	1211	1165
		Polynomial	1205	1159
	4	Linear	1220	1169
		Square Root	1219	1165
		Polynomial	1218	1164
FA	3	Linear	1221	1177
		Square Root	1209	1162
		Polynomial	1213	1167

4	Linear	1211	1160
	Square Root	1226	1173
	Polynomial	1238	1184

Supplementary Table 1- BIC and AIC values for the combination of the variable used in mixture argument and time function in three and four latent classes.

Chapter 5

Discussion

This thesis presents high accuracy of statistical approaches to: i) identify brain tissue at high risk of infarction in patients with stroke; and, ii) predict disability evolution in patients with multiple sclerosis.

These statistical approaches that are performed for stroke and multiple sclerosis patients had also several clinical interests. First, identifying the tissue at high risk (at one month) on the basis of the imaging data may guide clinicians in their time-critical decision process regarding the use of the tPA treatment. Second, clustering patients with MD and the predicting the evolution of disability may help the clinician in choosing the most adapted therapy regarding the choice of treatment.

In Chapter 2, the applicability of the machine learning methods to voxel-based multimodal MRI data was proved in identifying the tissue at high risk of infarction and also to classify tissues as healthy or infarcted in patients with stroke. We used eight advanced imaging parameters of diffusion-weighted imaging (DWI) and perfusion-weighted imaging (PWI) from 55 ischemic stroke patients. The classification methods (SVM, ANN, RF, ADA, and LR) showed high accuracies ($AUC_{roc} > 0.77$ for all methods). However, these classification methods were mainly compared regarding AUC_{pr} and there was no statistically significant difference between their performances.

In Chapter 3, we used an unsupervised method to cluster the multiple sclerosis patients on the basis of clinical and imaging trajectories. The study is the first that uses clustering of longitudinal data obtained over five years. We used the EDSS score, the lesion load, and fractional anisotropy trajectories to cluster the patients. The clusters showed three distinct forms of disease evolution: severe, stable, and moderate evolution. Moreover, the clusters which show the severe and stable evolution begin at EDSS 4, the limit of full ambulation without aid. This means that the unsupervised method was able to identify the patients who may have a critical evolution after EDSS score 4. Moreover, disease duration, GMV, and LL at study onset were significantly lower in the patients in the moderate cluster. These variables may be the prognostic factors for benign disease evolution over time. In addition, FA at study onset was significantly lower in the patients in the severe evolution cluster. Thus, a low FA value at study onset may be a prognostic factor for aggressive disease evolution.

In Chapter 4, we developed a model to predict disability evolution of multiple sclerosis patients. For this, we used the latent class linear mixed model to consider individual differences and unobserved groups among MS patients. The model used longitudinal clinical and imaging data. We established first a clinical model with clinical characteristics (age, disease duration, 25FW, and 9HPT scores). All clinical variables had significant effects on disease evolution and we obtained a high accuracy. When the gray matter volume, the lesion load, and the fractional atrophy were used to build the imaging model, the effect of the GMV was significantly different from zero. However, the clinical model had a better accuracy than the imaging model. Finally, the imaging parameters were added one by one, two by two, then three together to the clinical model in order to examine the importance of imaging data. However, the so combined models were not much better than the clinical model (as per the BIC criterion).

The latent class linear mixed model was also able to cluster the patients into three latent classes showing severe, stable and moderate evolution. However, the majority of patients were clustered in the stable class where there was no significant worsening or recovery

over five years.

As a consequence, given the high accuracy results, the models developed in this study are highly promising in predicting disability evolution of multiple sclerosis.

The limitation of our studies on multiple sclerosis is that we did not have access to the imaging parameters of the lesions. For example, the local lesion load and the location of the lesions might inform more about the risk of evolution toward severe disability. A future work may thus include the locally measured imaging parameters.

We hope that the statistical methods applied in this thesis will allow better integration and evaluation of future advances in medical imaging and statistical modeling. Finally, we hope to contribute to a better identification of stroke and multiple sclerosis patients for promising early-stage treatments.

Bibliography

- [1] Priorities in medicine. *The Lancet* 300, 7789 (1972), 1240. Originally published as Volume 2, Issue 7789.
- [2] ACHESON, E. D. Epidemiology of multiple sclerosis. *British Medical Bulletin* 33, 1 (1977), 9–14.
- [3] ACHIRON, A., AND BARAK, Y. Multiple sclerosis from probable to definite diagnosis: a 7-year prospective study. *Archives of Neurology* 57, 7 (2000), 974–979.
- [4] ALEXANDER, J. A., SHEPPARD, S., DAVIS, P. C., AND SALVERDA, P. Adult cerebrovascular disease: role of modified rapid fluid-attenuated inversion-recovery sequences. *American Journal of Neuroradiology* 17, 8 (1996), 1507–1513.
- [5] ARNOLD, D. Magnetic resonance spectroscopy: imaging axonal damage in MS. *Journal of Neuroimmunology* 98, 1 (1999), 2–6.
- [6] ARNOLD, D. L., MATTHEWS, P. M., FRANCIS, G., AND ANTEL, J. Proton magnetic resonance spectroscopy of human brain in vivo in the evaluation of multiple sclerosis: Assessment of the load of disease. *Magnetic Resonance in Medicine* 14, 1 (1990), 154–159.
- [7] ASTRUP, J., SIESJÖ, B. K., AND SYMON, L. Thresholds in cerebral ischemia - the ischemic penumbra. *Stroke* 12, 6 (1981), 723–725.
- [8] BAKSHI, R., BENEDICT, R., BERMEL, R., CARUTHERS, S., PULI, S., TJOA, C., FABIANO, A., AND JACOBS, L. T2 hypointensity in the deep gray matter of

- patients with multiple sclerosis: a quantitative magnetic resonance imaging study. *Arch Neurol.* 59, 1 (2002), 62–8.
- [9] BAKSHI, R., BENEDICT, R., BERMEL, R., AND JACOBS, L. Regional brain atrophy is associated with physical disability in multiple sclerosis: semiquantitative magnetic resonance imaging and relationship to clinical findings. *J Neuroimaging* 11, 2 (2001), 129–36.
- [10] BAMFORD, J., SANDERCOCK, P., DENNIS, M., WARLOW, C., AND BURN, J. Classification and natural history of clinically identifiable subtypes of cerebral infarction. *The Lancet* 337, 8756 (1991), 1521 – 1526.
- [11] BARCELLOS, L. F., KAMDAR, B. B., RAMSAY, P. P., DELOA, C., LINCOLN, R. R., CAILLIER, S., SCHMIDT, S., HAINES, J. L., PERICAK-VANCE, M. A., OKSENBERG, J. R., AND HAUSER, S. L. Clustering of autoimmune diseases in families with a high-risk for multiple sclerosis: a descriptive study. *The Lancet Neurology* 5, 11 (2006), 924 – 931.
- [12] BARKHOF, F., ELTON, M., AND ET AL. Functional correlates of callosal atrophy in relapsing-remitting multiple sclerosis patients. a preliminary MRI study. *J Neurol.* 245, 3 (1998), 153–8.
- [13] BARKHOF, F., SCHELTENS, P., AND ET AL. Relapsing-remitting multiple sclerosis: sequential enhanced MR imaging vs clinical findings in determining disease activity. *American Journal of Roentgenology* 159, 5.
- [14] BARNARD RO, T. M. Corpus callosum in multiple sclerosis. *Journal of Neurology, Neurosurgery, and Psychiatry* 37, 11 (1974), 1259–1264.
- [15] BAUER, D. J., AND CURRAN, P. J. Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods* 8, 3 (2003), 338–363.

- [16] BERGAMASCHI, R., BERZUINI, C., ROMANI, A., AND COSI, V. Predicting secondary progression in relapsing–remitting multiple sclerosis: a Bayesian analysis. *Journal of the Neurological Sciences* 189, 1 (2001), 13 – 21.
- [17] BERMEL, R., PULI, S., RUDICK, R., AND ET AL. Prediction of longitudinal brain atrophy in multiple sclerosis by gray matter magnetic resonance imaging T2 hypointensity. *Archives of Neurology* 62, 9 (2005), 1371–1376.
- [18] BITSCH, A., BRUHN, H., VOUGIOUKAS, V., STRINGARIS, A., LASSMANN, H., FRAHM, J., AND BRÜCK, W. Inflammatory CNS demyelination: Histopathologic correlation with in vivo quantitative proton MR spectroscopy. *American Journal of Neuroradiology* 20, 9 (1999), 1619–1627.
- [19] BJARTMAR, C., KIDD, G., AND ET AL. Neurological disability correlates with spinal cord axonal loss and reduced N-acetyl aspartate in chronic multiple sclerosis patients. *Ann Neurol* 48, 6 (2000), 893–901.
- [20] BOZDOGAN, H. Model selection and Akaike’s Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika* 52, 3 (Sep 1987), 345–370.
- [21] BRANT-ZAWADZKI, M., ATKINSON, D., DETRICK, M., BRADLEY, W. G., AND SCIDMORE, G. Fluid-attenuated inversion recovery (FLAIR) for assessment of cerebral infarction. *Stroke* 27, 7 (1996), 1187–1191.
- [22] BRÜCK, W., BITSCH, A., KOLENDA, H., BRÜCK, Y., STIEFEL, M., AND LASSMANN, H. Inflammatory central nervous system demyelination: Correlation of magnetic resonance imaging findings with lesion pathology. *Annals of Neurology* 42, 5 (1997), 783–793.
- [23] BREIMAN, L. Random forests. *Mach. Learn.* 45, 1 (Oct. 2001), 5–32.
- [24] BREIMAN, L., FRIEDMAN, J., OLSHEN, R., AND STONE, C. *Classification and Regression Trees*. Wadsworth and Brooks, Monterey, CA, 1984.

- [25] BREX, P. A., CICCARELLI, O., O'RIORDAN, J. I., SAILER, M., THOMPSON, A. J., AND MILLER, D. H. A longitudinal study of abnormalities on MRI and disability from multiple sclerosis. *New England Journal of Medicine* 346, 3 (2002), 158–164.
- [26] BROWNE, P., CHANDRARATNA, D., ANGOOD, C., TREMLETT, H., BAKER, C., TAYLOR, B. V., AND THOMPSON, A. J. Atlas of Multiple Sclerosis 2013: A growing global problem with widespread inequity. *Neurology* 83, 11 (Sept. 2014), 1022–1024.
- [27] BUTCHER, K., PARSONS, M., BAIRD, T., BARBER, A., DONNAN, G., DESMOND, P., TRESS, B., AND DAVIS, S. Perfusion thresholds in acute stroke thrombolysis. *Stroke* 34, 9 (2003), 2159–2164.
- [28] BUTTINELLI, C., CLEMENZI, A., BORRIELLO, G., DENARO, F., POZZILLI, C., AND FIESCHI, C. Mitoxantrone treatment in multiple sclerosis: a 5-year clinical and MRI follow-up. *European Journal of Neurology* 14, 11 (2007), 1281–1287.
- [29] CALIŃSKI, T., AND HARABASZ, J. A dendrite method for cluster analysis. *Communications in Statistics-Simulation and Computation* 3, 1 (1974), 1–27.
- [30] CAPLAN, L. Caplan's stroke: a clinical approach. *3rd Ed. Woburn: 2000* (2000).
- [31] CARAMANOS, Z., NARAYANAN, S., AND ARNOLD, D. L. 1H-MRS quantification of tNA and tCr in patients with multiple sclerosis: a meta-analytic review. *Brain* 128, 11 (2005), 2483–2506.
- [32] CARSWELL, R. *Pathological Anatomy. Illustrations of the Elementary Forms of Disease*. London: Longman, Orme, Brown, Green and Longman, 1838.
- [33] CERCIGNANI, M., INGLESE, M., PAGANI, E., COMI, G., AND FILIPPI, M. Mean diffusivity and fractional anisotropy histograms of patients with multiple sclerosis. *American Journal of Neuroradiology* 22, 5 (2001), 952–958.

- [34] CHOBANIAN, A. V., BAKRIS, G. L., BLACK, H. R., CUSHMAN, W. C., GREEN, L. A., IZZO, J. L., JONES, D. W., MATERSON, B. J., OPARIL, S., WRIGHT, J. T., ROCCELLA, EDWARD, J., AND ET AL. Seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure. 1206–1252.
- [35] CLERICO, M., CONTESSA, G., AND DURELLI, L. Interferon-beta1a for the treatment of multiple sclerosis. *Expert Opinion on Biological Therapy* 7, 4 (2007), 535–542.
- [36] COLES, A., WING, M., MOLYNEUX, P., PAOLILLO, A., AND ET AL. Monoclonal antibody treatment exposes three mechanisms underlying the clinical course of multiple sclerosis. *Ann Neurol.* 46, 3 (1999), 296–304.
- [37] COMMENGES, D., AND JACQMIN-GADDA, H. *Modèles biostatistiques pour l'épidémiologie.* de Boeck, Oct. 2015.
- [38] COMPSTON, A., AND COLES, A. Multiple sclerosis. *The Lancet* 359, 9313 (2002), 1221 – 1231.
- [39] CONFAVREUX, C., AND VUKUSIC, S. Natural history of multiple sclerosis: a unifying concept. *Brain* 129, 3 (2006), 606–616.
- [40] CONFAVREUX, C., VUKUSIC, S., AND ADELEINE, P. Early clinical predictors and progression of irreversible disability in multiple sclerosis: an amnesic process. *Brain* 126, 4 (2003), 770–782.
- [41] CONNOR, M. D., WALKER, R., MODI, G., AND WARLOW, C. P. Burden of stroke in black populations in sub-saharan africa. *The Lancet Neurology* 6, 3 (2007), 269 – 278.
- [42] CORTES, C., AND VAPNIK, V. Support-vector networks. In *Machine Learning* (1995), pp. 273–297.

- [43] COULTRAP, S., VEST, R., AND ET AL. Ca-MKII in cerebral ischemia. *Acta Pharmacol Sin.* 32, 7 (2011), 861–72.
- [44] DAMASCENO, A., AND ET AL. Prognostic indicators for long-term disability in multiple sclerosis patients. *Journal of the Neurological Sciences* 324, 1 (2013), 29 – 33.
- [45] DANIEL, K., WOLFE, C. D., BUSCH, M. A., AND MCKEVITT, C. What are the social consequences of stroke for working-aged adults? *Stroke* 40, 6 (2009), e431–e440.
- [46] DASGUPTA, R., AND FOWLER, C. J. Bladder, bowel and sexual dysfunction in multiple sclerosis. *Drugs* 63, 2 (Jan 2003), 153–166.
- [47] DAVIES, D. L., AND BOULDIN, D. W. A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* 1, 2 (Feb. 1979), 224–227.
- [48] DAVIS, J., AND GOADRICH, M. The relationship between precision-recall and ROC curves. In *Proceedings of the 23rd International Conference on Machine Learning* (New York, NY, USA, 2006), ICML '06, ACM, pp. 233–240.
- [49] DE STEFANO, N., NARAYANAN, S., FRANCIS, G., AND ET AL. Evidence of axonal damage in the early stages of multiple sclerosis and its relevance to disability. *Archives of Neurology* 58, 1 (2001), 65–70.
- [50] DEGENHARDT, A., RAMAGOPALAN, S., AND ET AL. Clinical prognostic factors in multiple sclerosis: a natural history review. *Nat Rev Neurol.* 5, 12 (2009), 672–82.
- [51] DEMPSTER, A. P., LAIRD, N. M., AND RUBIN, D. B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* 39, 1 (1977), 1–38.
- [52] DIRNAGL, U., IADECOLA, C., AND MOSKOWITZ, M. A. Pathobiology of ischaemic stroke: an integrated view. *Trends in Neurosciences* 22, 9 (1999), 391 – 397.

- [53] DONNAN, G. A., FISHER, M., MACLEOD, M., AND DAVIS, S. M. Stroke. *The Lancet* 371, 9624 (2008), 1612 – 1623.
- [54] EBERS, G. Optic neuritis and multiple sclerosis. *Archives of Neurology* 42, 7 (1985), 702–704.
- [55] EBERS, G. C., BULMAN, D. E., SADOVNICK, A. D., PATY, D. W., WARREN, S., HADER, W., MURRAY, T. J., SELAND, T. P., DUQUETTE, P., GREY, T., NELSON, R., NICOLLE, M., AND BRUNET, D. A population-based study of multiple sclerosis in twins. *New England Journal of Medicine* 315, 26 (1986), 1638–1642.
- [56] EMILIA SBARDELLA, FRANCESCA TONA, N. P., AND PANTANO, P. DTI measurements in multiple sclerosis: Evaluation of brain damage and clinical implications. *Multiple Sclerosis International 2013* (2013), 11 pages.
- [57] ESHAGHI, A., BODINI, B., RIDGWAY, G. R., GARCÍA-LORENZO, D., TOZER, D. J., SAHRAIAN, M. A., THOMPSON, A. J., AND CICCARELLI, O. Temporal and spatial evolution of grey matter atrophy in primary progressive multiple sclerosis. *NeuroImage* 86, Supplement C (2014), 257 – 264.
- [58] EVERITT, B., LANDAU, S., LEESE, M., AND STAHL, D. *Cluster analysis*, 5th ed. Wiley, 2011.
- [59] FEIGIN, V., MENSAH, G. L., AND ET AL. Atlas of the global burden of stroke (1990-2013): The GBD 2013 study. *Neuroepidemiology* 45, 9624 (2015), 230–236.
- [60] FILIPPI, C. G., ANDREWS, T., GONYEA, J. V., LINNELL, G., AND CAULEY, K. A. Magnetic resonance diffusion tensor imaging and tractography of the lower spinal cord: application to diastematomyelia and tethered cord. *European Radiology* 20, 9 (Sep 2010), 2194–2199.
- [61] FILIPPI, M. The role of non-conventional magnetic resonance techniques in monitoring evolution of multiple sclerosis. *Journal of Neurology, Neurosurgery, and Psychiatry* 64 Suppl 1 (May 1998), S52—8.

- [62] FILIPPI, M. Non-conventional MR techniques to monitor the evolution of multiple sclerosis. *Neurological Sciences* 22, 2 (Apr 2001), 195–200.
- [63] FILIPPI, M., CERCIGNANI, M., INGLESE, M., HORSFIELD, M., AND COMI, G. Diffusion tensor magnetic resonance imaging in multiple sclerosis. *Neurology* 56, 3 (2001), 304–311.
- [64] FILIPPI, M., AND ET AL. MRI criteria for the diagnosis of multiple sclerosis: Magnims consensus guidelines. *The Lancet Neurology* 15, 3 (2016), 292 – 303.
- [65] FILIPPI, M., AND ROCCA, M. A. Magnetization transfer magnetic resonance imaging of the brain, spinal cord, and optic nerve. *Neurotherapeutics* 4, 3 (Jul 2007), 401–413.
- [66] FILIPPI, M., ROCCA, M. A., BARKHOF, F., BRÜCK, W., CHEN, J. T., COMI, G., DELUCA, G., STEFANO, N. D., ERICKSON, B. J., EVANGELOU, N., FAZEKAS, F., GEURTS, J. J., LUCCHINETTI, C., MILLER, D. H., PELLETIER, D., POPESCU, B. F. G., AND LASSMANN, H. Association between pathological and MRI findings in multiple sclerosis. *The Lancet Neurology* 11, 4 (2012), 349 – 360.
- [67] FISNIKU, L. K., CHARD, D. T., JACKSON, J. S., ANDERSON, V. M., ALTMANN, D. R., MISZKIEL, K. A., THOMPSON, A. J., AND MILLER, D. H. Gray matter atrophy is related to long-term disability in multiple sclerosis. *Annals of Neurology* 64, 3 (2008), 247–254.
- [68] FLETCHER, R. *Newton-Like Methods*. John Wiley and Sons, Ltd, 2000, pp. 44–79.
- [69] FREEDMAN, M., BAR-OR, A., OGER, J., TRABOULSEE, A., PATRY, D., YOUNG, C., OLSSON, T., LI, D., HARTUNG, H.-P., KRANTZ, M., FERENCZI, L., AND VERCO, T. A phase III study evaluating the efficacy and safety of MBP8298 in secondary progressive MS. *Neurology* 77, 16 (2011), 1551–1560.
- [70] FREEDMAN, M. S., COMI, G., STEFANO, N. D., BARKHOF, F., POLMAN, C. H., UITDEHAAG, B. M., LEHR, L., STUBINSKI, B., AND KAPPOS, L. Moving toward

- earlier treatment of multiple sclerosis: Findings from a decade of clinical trials and implications for clinical practice. *Multiple Sclerosis and Related Disorders* 3, 2 (2014), 147 – 155.
- [71] FROMONT, A., BINQUET, C., SAULEAU, E. A., FOURNEL, I., BELLISARIO, A., ADNET, J., WEILL, A., VUKUSIC, S., CONFAVREUX, C., DEBOUVERIE, M., CLERC, L., BONITHON-KOPP, C., AND MOREAU, T. Geographic variations of multiple sclerosis in France. *Brain* 133, 7 (2010), 1889–1899.
- [72] FU, L., MATTHEWS, P., AND ET AL. Imaging axonal damage of normal-appearing white matter in multiple sclerosis. *Brain* 121, Pt 1 (1998), 103–13.
- [73] GALEA, I., WARD-ABEL, N., AND HEESEN, C. Relapse in multiple sclerosis. *BMJ* 350 (2015).
- [74] GE, Y. Multiple sclerosis: The role of MR imaging. *American Journal of Neuro-radiology* 27, 6 (2006), 1165–1176.
- [75] GE, Y., LAW, M., AND GROSSMAN, R. I. Applications of diffusion tensor MR imaging in multiple sclerosis. *Annals of the New York Academy of Sciences* 1064, 1 (2005), 202–219.
- [76] GE, Y., LAW, M., JOHNSON, G., HERBERT, J., BABB, J. S., MANNON, L. J., AND GROSSMAN, R. I. Preferential occult injury of corpus callosum in multiple sclerosis measured by diffusion tensor imaging. *Journal of Magnetic Resonance Imaging* 20, 1 (2004), 1–7.
- [77] GEAN-MARTON, A. D., VEZINA, L. G., MARTON, K. I., STIMAC, G. K., PEYSTER, R. G., TAVERAS, J. M., AND DAVIS, K. R. Abnormal corpus callosum: a sensitive and specific indicator of multiple sclerosis. *Radiology* 180, 1 (1991), 215–221.
- [78] GENOLINI, C., ALACOQUE, X., SENTENAC, M., AND ARNAUD, C. kml and kml3d: R packages to cluster longitudinal data. *Journal of Statistical Software* 65, 4 (2015), 1–34.

- [79] GIRI, S., CHUNG, Y.-C., MERCHANT, A., MIHAI, G., RAJAGOPALAN, S., RAMAN, S. V., AND SIMONETTI, O. P. T2 quantification for improved detection of myocardial edema. *Journal of Cardiovascular Magnetic Resonance* 11, 1 (Dec 2009), 56.
- [80] GOLDENBERG, M. Multiple sclerosis review. *Pharmacy and Therapeutics* 37, 3 (2012), 175–184.
- [81] GUERRERO, W. R., OKUN, M. S., AND MCFARLAND, N. R. Encephalopathy, hypoglycemia, and flailing extremities: A case of bilateral chorea–ballism associated with diabetic ketoacidosis. In *Tremor and other hyperkinetic movements* (2012).
- [82] GURNEY, K. *An Introduction to Neural Networks*, vol. ISBN:1857286731. 1997.
- [83] HACKE, W., KASTE, M., BLUHMKI, E., BROZMAN, M., DÁVALOS, A., GUIDETTI, D., LARRUE, V., LEES, K. R., MEDEGHRI, Z., MACHNIG, T., SCHNEIDER, D., VON KUMMER, R., WAHLGREN, N., AND TONI, D. Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. *New England Journal of Medicine* 359, 13 (2008), 1317–1329.
- [84] HANLEY, J. Receiver operating characteristic (ROC) methodology: the state of the art. *Crit Rev Diagn Imaging* 29, 3 (1989), 307–35.
- [85] HANNOUN, S., BAGORY, M., DURAND-DUBIEF, F., IBARROLA, D., COMTE, J.-C., CONFAVREUX, C., COTTON, F., AND SAPPEY-MARINIER, D. Correlation of diffusion and metabolic alterations in different clinical forms of multiple sclerosis. *PLOS ONE* 7, 3 (03 2012), 1–7.
- [86] HARTUNG, H.-P., AND ET AL. Mitoxantrone in progressive multiple sclerosis: a placebo-controlled, double-blind, randomised, multicentre trial. *The Lancet* 360 (2002), 2018 – 2025.
- [87] HASAN, K. M., GUPTA, R. K., SANTOS, R. M., WOLINSKY, J. S., AND NARAYANA, P. A. Diffusion tensor fractional anisotropy of the normal-appearing

- seven segments of the corpus callosum in healthy adults and relapsing-remitting multiple sclerosis patients. *Journal of Magnetic Resonance Imaging* 21, 6 (2005), 735–743.
- [88] HAWKER, K. Progressive Multiple Sclerosis: Characteristics and Management. *Neurologic Clinics* 29 (2011), 423 – 434.
- [89] HAWKINS, D. S., ALLEN, D. M., AND STROMBERG, A. J. Determining the number of components in mixtures of linear models. *Computational Statistics and Data Analysis* 38, 1 (2001), 15–48.
- [90] HECHT-NIELSEN, R. Theory of the backpropagation neural network. In *International 1989 Joint Conference on Neural Networks* (1989), pp. 593–605 vol.1.
- [91] HJORT, N., CHRISTENSEN, S., SØLLING, C., ASHKANIAN, M., WU, O., RØHL, L., GYLDENSTED, C., ANDERSEN, G., AND ØSTERGAARD, L. Ischemic injury detected by diffusion imaging 11 minutes after stroke. *Annals of Neurology* 58, 3 (2005), 462–465.
- [92] HOLMØY, T. Immunopathogenesis of multiple sclerosis: concepts and controversies. *Acta Neurologica Scandinavica* 115 (2007), 39–45.
- [93] HORSFIELD MA, LARSSON HB, J. D. G. A. Diffusion magnetic resonance imaging in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 64, Suppl 1 (1998), 80–4.
- [94] HOWARD, V., CUSHMAN, M. J., PULLEY, L., GOMEZ, C., GO, R., PRINEAS, C. R., GRAHAM, A. J., MOY, C. S., AND HOWARD, G. The reasons for geographic and racial differences in stroke study: Objectives and design. *Neuroepidemiology* 25 (2005), 135–143.
- [95] IADECOLA, C., AND GORELICK, P. B. Hypertension, angiotensin, and stroke: beyond blood pressure. *Stroke* 35, 2 (2004), 348–350.
- [96] INGLESE, M., AND BESTER, M. Diffusion imaging in multiple sclerosis: research and clinical implications. *NMR in Biomedicine* 23, 7 (2010), 865–872.

- [97] ION-MĂRGINEANU, A., KOCEVAR, G., STAMILE, C., SIMA, D. M., DURAND-DUBIEF, F., VAN HUFFEL, S., AND SAPPEY-MARINIER, D. Machine learning approach for classifying multiple sclerosis courses by combining clinical data with lesion loads and magnetic resonance metabolic features. *Frontiers in Neuroscience* 11 (2017), 398.
- [98] ISLAM, T., GAUDERMAN, W. J., COZEN, W., AND MACK, T. M. Childhood sun exposure influences risk of multiple sclerosis in monozygotic twins. *Neurology* 69, 4 (2007), 381–388.
- [99] JACK, C., PETERSEN, R., XU, Y., AND ET AL. Medial temporal atrophy on MRI in normal aging and very mild alzheimer’s disease. *Neurology* 49, 3 (1997), 786–794.
- [100] JACOBS, L. D., COOKFAIR, D. L., RUDICK, R. A., HERNDON, R. M., RICHERT, J. R., SALAZAR, A. M., FISCHER, J. S., GOODKIN, D. E., GRANGER, C. V., SIMON, J. H., ALAM, J. J., BARTOSZAK, D. M., BOURDETTE, D. N., BRAIMAN, J., BROWNSCHIEDLE, C. M., COATS, M. E., COHAN, S. L., DOUGHERTY, D. S., KINKEL, R. P., MASS, M. K., MUNSCHAUER, F. E., PRIORE, R. L., PULLICINO, P. M., SCHEROKMAN, B. J., WEINSTOCK-GUTTMAN, B., AND WHITHAM, R. H. Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. *Annals of Neurology* 39, 3 (1996), 285–294.
- [101] JAIN, A. K., MAO, J., AND MOHIUDDIN, K. Artificial neural networks: A tutorial. *IEEE Computer* 29 (1996), 31–44.
- [102] JAMES A. FREEMAN, D. M. S. *Neural networks: algorithms, applications, and programming techniques*, vol. ISBN:0-201-51376-5. 1991.
- [103] J. CRUVEILHIER. *Anatomie pathologique du corps humain, ou, Description avec figures lithographiées et coloriées, des diverses altérations morbides dont le corps humain est susceptible*. Paris. Chez Baillière, 1842.
- [104] JF., K. Geography in multiple sclerosis. *J Neurol.* 215, 1 (1977), 1–26.

- [105] JF., K. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 33, 11 (1983), 1444–52.
- [106] JOHNSON, K., BROOKS, B., COHEN, J., FORD, C., GOLDSTEIN, J., LISAK, R., MYERS, L., PANITCH, H., ROSE, J., AND SCHIFFER, R. Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: results of a phase III multicenter, double-blind placebo-controlled trial. The Copolymer 1 Multiple Sclerosis Study Group. *Neurology* 45, 7 (1995), 1268–76.
- [107] JOHNSTON, S. C., MENDIS, S., AND MATHERS, C. D. Global variation in stroke burden and mortality: estimates from monitoring, surveillance, and modelling. *The Lancet Neurology* 8, 4 (2009), 345 – 354.
- [108] KAMPMAN, M. T., WILSGAARD, T., AND MELLGREN, S. I. Outdoor activities and diet in childhood and adolescence relate to MS risk above the arctic circle. *Journal of Neurology* 254, 4 (Apr 2007), 471–477.
- [109] KANDEL, E. R. The molecular biology of memory: cAMP, PKA, CRE, CREB-1, CREB-2, and CPEB. *Molecular Brain* 5, 1 (May 2012), 14.
- [110] KAPPOS, L., AND ET AL. Effect of early versus delayed interferon beta-1b treatment on disability after a first clinical event suggestive of multiple sclerosis: a 3-year follow-up analysis of the BENEFIT study. *The Lancet* 370, 9585 (2007), 389 – 397.
- [111] KARATZOGLOU, A., MEYER, D., AND HORNIK, K. Support vector machines in R. *Journal of Statistical Software, Articles* 15, 9 (2006), 1–28.
- [112] KATZ, D., TAUBENBERGER, J. K., CANNELLA, B., MCFARLIN, D. E., RAINE, C. S., AND MCFARLAND, H. F. Correlation between magnetic resonance imaging findings and lesion development in chronic, active multiple sclerosis. *Annals of Neurology* 34, 5 (1993), 661–669.
- [113] KAUFMAN, L., AND ROUSSEEUW, P. J. *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley, 1990.

- [114] KIDWELL, C. S., ALGER, J. R., AND SAVER, J. L. Beyond mismatch. 2729–2735.
- [115] KOCEVAR, G., STAMILE, C., HANNOUN, S., COTTON, F., VUKUSIC, S., DURAND-DUBIEF, F., AND SAPPEY-MARINIER, D. Graph theory-based brain connectivity for automatic classification of multiple sclerosis clinical courses. *Frontiers in Neuroscience* 10 (2016), 478.
- [116] KOMÁREK, A. A new R package for bayesian estimation of multivariate normal mixtures allowing for selection of the number of components and interval-censored data. *Computational Statistics and Data Analysis* 53, 12 (2009), 3932 – 3947.
- [117] KRYSZCZUK, K., AND HURLEY, P. *Estimation of the Number of Clusters Using Multiple Clustering Validity Indices*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010, pp. 114–123.
- [118] KURTZKE, J. F. Multiple sclerosis in time and space - geographic clues to cause. *Journal of NeuroVirology* 6, Suppl 2 (2000), 134 –140.
- [119] LA., R. Multiple sclerosis: It’s not the disease you thought it was. *Clinical Medicine and Research* 1, 1 (2003), 57–60.
- [120] LAIRD, N. M., AND WARE, J. H. Random-effects models for longitudinal data. *Biometrics* 38, 4 (1982), 963–974.
- [121] LANSBERG, M., O’BRIEN, M., TONG, D., MOSELEY, M., AND ALBERS, G. Evolution of cerebral infarct volume assessed by diffusion-weighted magnetic resonance imaging. *Archives of Neurology* 58, 4 (2001), 613–617.
- [122] LARSSON, H. B. W., CHRISTIANSEN, P., JENSEN, M., FREDERIKSEN, J., HELTBERG, A., OLESEN, J., AND HENRIKSEN, O. Localized in vivo proton spectroscopy in the brain of patients with multiple sclerosis. *Magnetic Resonance in Medicine* 22, 1 (1991), 23–31.
- [123] LASSMANN, H., BRÜCK, W., AND LUCCHINETTI, C. F. The immunopathology of multiple sclerosis: An overview. *Brain Pathology* 17, 2 (2007), 210–218.

- [124] LAWTON, M., AND ET AL. A longitudinal model for disease progression was developed and applied to multiple sclerosis. *Journal of Clinical Epidemiology* 68, 11 (2015), 1355 – 1365.
- [125] LEIVA-SALINAS C, W. M. Imaging of ischemic stroke. *Neuroimaging clinics of North America* 20, 4.
- [126] LINDSTROM, M. J., AND BATES, D. M. Newton-raphson and em algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association* 83, 404 (1988), 1014–1022.
- [127] LLUFRIU, S., BLANCO, Y., MARTINEZ-HERAS, E., CASANOVA-MOLLA, J., GABILONDO, I., SEPULVEDA, M., FALCON, C., BERENGUER, J., BARGALLO, N., VILLOSLADA, P., GRAUS, F., VALLS-SOLE, J., AND SAIZ, A. Influence of corpus callosum damage on cognition and physical disability in multiple sclerosis: A multimodal study. *PLOS ONE* 7, 5 (05 2012), 1–7.
- [128] LOPEZ, A. D., MATHERS, C. D., EZZATI, M., JAMISON, D. T., AND MURRAY, C. J. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *The Lancet* 367, 9524 (2006), 1747 – 1757.
- [129] LUBLIN, F. D. New multiple sclerosis phenotypic classification. *Eur Neurol* 72, suppl 1 (2014), 1–5.
- [130] LUBLIN, F. D., REINGOLD, S. C., AND ON CLINICAL TRIALS OF NEW AGENTS IN MULTIPLE SCLEROSIS, N. M. S. S. U. A. C. Defining the clinical course of multiple sclerosis: Results of an international survey. *Neurology* 46, 4 (1996), 907–911.
- [131] LUCAS, R., PONSONBY, A.-L., DEAR, K., VALERY, P., PENDER, M., TAYLOR, B., KILPATRICK, T., DWYER, T., COULTHARD, A., CHAPMAN, C., VAN DER MEI, I., WILLIAMS, D., AND MCMICHAEL, A. Sun exposure and vitamin D are independent risk factors for CNS demyelination. *Neurology* 76, 6 (2011), 540–548.

- [132] LYDEN, P., AND ZIVIN, J. Hemorrhagic transformation after cerebral ischemia: mechanisms and incidence. *Cerebrovasc Brain Metab Rev.* 5, 1 (1993), 1–16.
- [133] MACQUEEN, J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics* (Berkeley, Calif., 1967), University of California Press, pp. 281–297.
- [134] MARQUARDT, D. W. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics* 11, 2 (1963), 431–441.
- [135] MARRIE, R. A. Environmental risk factors in multiple sclerosis aetiology. *The Lancet Neurology* 3, 12 (2004), 709 – 718.
- [136] McDONALD, W. I., COMPSTON, A., EDAN, G., GOODKIN, D., HARTUNG, H.-P., LUBLIN, F. D., MCFARLAND, H. F., PATY, D. W., POLMAN, C. H., REINGOLD, S. C., SANDBERG-WOLLHEIM, M., SIBLEY, W., THOMPSON, A., VAN DEN NOORT, S., WEINSHENKER, B. Y., AND WOLINSKY, J. S. Recommended diagnostic criteria for multiple sclerosis: Guidelines from the international panel on the diagnosis of multiple sclerosis. *Annals of Neurology* 50, 1 (2001), 121–127.
- [137] MCLACHLAN, G. J., AND PEEL, D. *Finite mixture models*. Wiley Series in Probability and Statistics, New York, 2000.
- [138] MENDES A, S. M. Classical immunomodulatory therapy in multiple sclerosis: how it acts, how it works. *Arquivos de Neuro-Psiquiatria* 69 (06 2011), 536 – 543.
- [139] MEULMAN, T. J., BAKKER, J., AND VAN DEN BOS, E. J. Ischemic cardiomyopathy and cerebral infarction in a young patient associated with khat chewing. In *Case reports in radiology* (2015).
- [140] MEYER, D., AND WIEN, T. U. Support vector machines. the interface to libsvm in package e1071. online-documentation of the package e1071 for R.

- [141] MILLEFIORINI, E., GASPERINI, C., POZZILLI, C., D'ANDREA, F., BASTIANELLO, S., TROJANO, M., MORINO, S., MORRA, V., BOZZAO, A., CALO', A., BERNINI, M., GAMBI, D., AND PRENCIPE, M. Randomized placebo-controlled trial of mitoxantrone in relapsing-remitting multiple sclerosis: 24-month clinical and MRI outcome. *J Neurol.* 244, 3 (1997), 153–9.
- [142] MILLER, D., GROSSMAN, R., REINGOLD, S., AND MCFARLAND, H. The role of magnetic resonance techniques in understanding and managing multiple sclerosis. *Brain* 121, Pt 1 (1998), 3–24.
- [143] MILLER, D. H., AND ET AL. Clinically isolated syndromes. *The Lancet Neurology* 11, 2 (2012), 157 – 169.
- [144] MILLIGAN, G. W., AND COOPER, M. C. An examination of procedures for determining the number of clusters in a data set. *Psychometrika* 50, 2 (Jun 1985), 159–179.
- [145] MILLS, R., AND YOUNG, C. A medical definition of fatigue in multiple sclerosis. *QJM: An International Journal of Medicine* 101, 1 (2008), 49–60.
- [146] MILO, R., AND KAHANA, E. Multiple sclerosis: Geoepidemiology, genetics and the environment. *Autoimmunity Reviews* 9, 5 (2010), A387 – A394. Special Issue on The Environment Geoepidemiology and Autoimmune Diseases.
- [147] MUNGER, K., LEVIN, L., HOLLIS, B., HOWARD, N., AND ASCHERIO, A. Serum 25-hydroxyvitamin d levels and risk of multiple sclerosis. *JAMA* 296, 23 (2006), 2832–2838.
- [148] MUTHEN, B. O., AND MUTHEN, L. K. *Mplus (Version 4.1) [Computer software]*, 1998–2006).
- [149] MUTHÉN, B., AND SHEDDEN, K. Finite mixture modeling with mixture outcomes using the em algorithm. *Biometrics* 55, 2 (1999), 463–469.

- [150] MUTHURAMAN, M., FLEISCHER, V., KOLBER, P., LUESSI, F., ZIPP, F., AND GROPPA, S. Structural brain network characteristics can differentiate CIS from early RRMS. *Frontiers in Neuroscience* 10 (2016), 14.
- [151] NOSEWORTHY, J. H., LUCCHINETTI, C., RODRIGUEZ, M., AND WEINSHENKER, B. G. Multiple sclerosis. *New England Journal of Medicine* 343, 13 (2000), 938–952.
- [152] OLERUP, O., AND HILLERT, J. HLA class II-associated genetic susceptibility in multiple sclerosis: A critical evaluation. *Tissue Antigens* 38, 2 (1991), 1–15.
- [153] OPARA J, JARACZ K, B. W. Quality of life in mulptiple sclerosis. *Journal of Medicine and Life* 3, 4 (2010), 352–358.
- [154] OWOLABI, M. O., UGOYA, S., AND PLATZ, T. Racial disparity in stroke risk factors: the berlin–ibadan experience; a retrospective study. *Acta Neurologica Scandinavica* 119, 2 (2009), 81–87.
- [155] OZENNE, B., SUBTIL, F., AND MAUCORT-BOULCH, D. The precision–recall curve overcame the optimism of the receiver operating characteristic curve in rare diseases. *Journal of Clinical Epidemiology* 68, 8 (2015), 855 – 859.
- [156] OZTURK, A., SMITH, S., GORDON-LIPKIN, E., HARRISON, D., SHIEE, N., PHAM, D., CAFFO, B., CALABRESI, P., AND REICH, D. MRI of the corpus callosum in multiple sclerosis: association with disability. *Multiple Sclerosis Journal* 16, 2 (2010), 166–177.
- [157] PAOLILLO, A., COLES, A., MOLYNEUX, P., GAWNE-CAIN, M., MACMANUS, D., BARKER, G., COMPSTON, D., AND MILLER, D. Quantitative MRI in patients with secondary progressive MS treated with monoclonal antibody Campath 1H. *Neurology* 53, 4 (1999), 751–7.
- [158] PARSONS, M. W., BARBER, P. A., CHALK, J., DARBY, D. G., ROSE, S., DESMOND, P. M., GERRATY, R. P., TRESS, B. M., WRIGHT, P. M., DONNAN, G. A., AND DAVIS, S. M. Diffusion- and perfusion-weighted MRI response to thrombolysis in stroke. *Annals of Neurology* 51, 1 (2002), 28–37.

- [159] PATY DW, L. D. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. II. MRI analysis results of a multicenter, randomized, double-blind, placebo-controlled trial. UBC MS/MRI Study Group and the IFNB Multiple Sclerosis Study Group. *Neurology* 43, 4 (1993), 662–7.
- [160] PIEHL, F. A changing treatment landscape for multiple sclerosis: challenges and opportunities. *Journal of Internal Medicine* 275, 4 (2014), 364–381.
- [161] POLMAN, C. H., REINGOLD, S. C., BANWELL, B., CLANET, M., COHEN, J. A., FILIPPI, M., FUJIHARA, K., HAVRDOVA, E., HUTCHINSON, M., KAPPOS, L., LUBLIN, F. D., MONTALBAN, X., O’CONNOR, P., SANDBERG-WOLLHEIM, M., THOMPSON, A. J., WAUBANT, E., WEINSHENKER, B., AND WOLINSKY, J. S. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of Neurology* 69, 2 (2011), 292–302.
- [162] POLMAN, C. H., REINGOLD, S. C., EDAN, G., FILIPPI, M., HARTUNG, H.-P., KAPPOS, L., LUBLIN, F. D., METZ, L. M., MCFARLAND, H. F., O’CONNOR, P. W., SANDBERG-WOLLHEIM, M., THOMPSON, A. J., WEINSHENKER, B. G., AND WOLINSKY, J. S. Diagnostic criteria for multiple sclerosis: 2005 revisions to the “McDonald criteria”. *Annals of Neurology* 58, 6 (2005), 840–846.
- [163] POSER, C. M., PATY, D. W., SCHEINBERG, L., MCDONALD, W. I., DAVIS, F. A., EBERS, G. C., JOHNSON, K. P., SIBLEY, W. A., SILBERBERG, D. H., AND TOURTELLOTTE, W. W. New diagnostic criteria for multiple sclerosis: Guidelines for research protocols. *Annals of Neurology* 13, 3 (1983), 227–231.
- [164] PREZIOSA, P., ROCCA, M. A., MESAROS, S., PAGANI, E., STOSIC-OPINCAL, T., KACAR, K., ABSINTA, M., CAPUTO, D., DRULOVIC, J., COMI, G., AND FILIPPI, M. Intrinsic damage to the major white matter tracts in patients with different clinical phenotypes of multiple sclerosis: A voxelwise diffusion-tensor MR study. *Radiology* 260, 2 (2011), 541–550.

- [165] PROUST, C., AND JACQMIN-GADDA, H. Estimation of linear mixed models with a mixture of distribution for the random effects. *Computer Methods and Programs in Biomedicine* 78, 2 (2005), 165 – 173.
- [166] PROUST-LIMA, C., PHILIPPS, V., AND LIQUET, B. Estimation of extended mixed models using latent classes and latent processes: The R package lamm. *Journal of Statistical Software, Articles* 78, 2 (2017), 1–56.
- [167] RAMAGOPALAN, S. V., DOBSON, R., MEIER, U. C., AND GIOVANNONI, G. Multiple sclerosis: risk factors, prodromes, and potential causal pathways. *The Lancet Neurology* 9, 7 (2010), 727 – 739.
- [168] RANA, S., ALBAYRAM, S., LIN, D. D., AND YOUSEM, D. M. Diffusion-weighted imaging and apparent diffusion coefficient maps in a case of intracerebral abscess with ventricular extension. *American Journal of Neuroradiology* 23, 1 (2002), 109–112.
- [169] RAO, S., LEO, G., BERNARDIN, L., AND UNVERZAGT, F. Cognitive dysfunction in multiple sclerosis. i. frequency, patterns, and prediction. *Neurology* 41, 5 (1991), 685–91.
- [170] RAY, S., AND TURI, R. H. Determination of number of clusters in k-means clustering and application in colour segmentation. In *The 4th International Conference on Advances in Pattern Recognition and Digital Techniques* (1999), pp. 137–143.
- [171] READ, S. J., HIRANO, T., ABBOTT, D. F., MARKUS, R., SACHINIDIS, J. I., TOCHON-DANGUY, H. J., CHAN, J. G., EGAN, G. F., SCOTT, A. M., BLADIN, C. F., MCKAY, W. J., AND DONNAN, G. A. The fate of hypoxic tissue on 18F-fluoromisonidazole positron emission tomography after ischemic stroke. *Annals of Neurology* 48, 2 (2000), 228–235.
- [172] RENARD, D., CASTELNOVO, G., BOUSQUET, P.-J., DE CHAMPFLEUR, N., DE SEZE, J., VERMERSCH, P., AND LABAUGE, P. Brain MRI findings in long-

- standing and disabling multiple sclerosis in 84 patients. *Clinical Neurology and Neurosurgery* 112, 4 (2010), 286 – 290.
- [173] RIMKUS, C. D. M., JUNQUEIRA, T. D. F., AND ET AL. Corpus Callosum Microstructural Changes Correlate with Cognitive Dysfunction in Early Stages of Relapsing-Remitting Multiple Sclerosis: Axial and Radial Diffusivities Approach. *Multiple Sclerosis International* 2011 (2011), 7 pages.
- [174] RINGLEB, P., SCHELLINGER, P., SCHRANZ, C., AND HACKE, W. Thrombolytic therapy within 3 to 6 hours after onset of ischemic stroke. *Stroke* 33, 5 (2002), 1437–1441.
- [175] ROEBUCK, J. R., HAKER, S. J., MITSOURAS, D., RYBICKI, F. J., TEMPANY, C. M., AND MULKERN, R. V. Carr-Purcell-Meiboom-Gill imaging of prostate cancer: quantitative T2 values for cancer discrimination. *Magnetic Resonance Imaging* 27, 4 (2009), 497 – 502.
- [176] ROSEN, B. R., BELLIVEAU, J. W., VEVEA, J. M., AND BRADY, T. J. Perfusion imaging with NMR contrast agents. *Magnetic Resonance in Medicine* 14, 2 (1990), 249–265.
- [177] ROTHER, J., SCHELLINGER, P., GASS, A., SIEBLER, M., VILLRINGER, A., FIEBACH, J., FIEHLER, J., JANSEN, O., KUCINSKI, T., SCHODER, V., SZABO, K., JUNGE-HÜLSING, G., HENNERICI, M., ZEUMER, H., SARTOR, K., WEILLER, C., AND HACKE, W. Effect of intravenous thrombolysis on MRI parameters and functional outcome in acute stroke. *Stroke* 33, 10 (2002), 2438–2445.
- [178] ROVARIS, M., GASS, A., BAMMER, R., HICKMAN, S. J., CICCARELLI, O., MILLER, D. H., AND FILIPPI, M. Diffusion MRI in multiple sclerosis. *Neurology* 65, 10 (2005), 1526–1532.
- [179] ROVIRA, L., AUGER, C., AND ALONSO, J. Magnetic resonance monitoring of lesion evolution in multiple sclerosis. *Therapeutic Advances in Neurological Disorders* 6, 5 (2013), 298–310.

- [180] RUDICK, R. A., CUTTER, G., AND REINGOLD, S. The multiple sclerosis functional composite: a new clinical outcome measure for multiple sclerosis trials. *Multiple Sclerosis Journal* 8, 5 (2002), 359–365.
- [181] RUEDA, F., HYGINO, L. C., DOMINGUES, R., VASCONCELOS, C., PAPAIS-ALVARENGA, R., AND GASPARETTO, E. Diffusion tensor MR imaging evaluation of the corpus callosum of patients with multiple sclerosis. *Arquivos de Neuro-Psiquiatria* 66 (09 2008), 449 – 453.
- [182] RUNMARKER, B., ANDERSSON, C., ODÉN, A., AND ANDERSEN, O. Prediction of outcome in multiple sclerosis based on multivariate models. *Journal of Neurology* 241, 10 (Oct 1994), 597–604.
- [183] SÃ, M. J. A. Physiopathology of symptoms and signs in multiple sclerosis. *Arquivos de Neuro-Psiquiatria* 70 (09 2012), 733 – 740.
- [184] SADOVNICK, A. D., ARMSTRONG, H., RICE, G. P. A., BULMAN, D., HASHIMOTO, L., PARTY, D. W., HASHIMOTO, S. A., WARREN, S., HADER, W., MURRAY, T. J., SELAND, T. P., METZ, L., BELL, R., DUQUETTE, P., GRAY, T., NELSON, R., WEINSHENKAR, B., BRUNT, D., AND EBERS, G. C. A population-based study of multiple sclerosis in twins: Update. *Annals of Neurology* 33, 3 (1993), 281–285.
- [185] SAHRAIAN, M. A., AND ESHAGHI, A. Role of MRI in diagnosis and treatment of multiple sclerosis. *Clinical Neurology and Neurosurgery* 112, 7 (2010), 609 – 615.
- [186] SAJJA, B. R., WOLINSKY, J. S., AND NARAYANA, P. A. Proton magnetic resonance spectroscopy in multiple sclerosis. *Neuroimaging Clinics of North America* 19, 1 (2009), 45 – 58. Multiple Sclerosis, Part II: Nonconventional MRI Techniques.
- [187] SCALFARI, A., NEUHAUS, A., DAUMER, M., MURARO, P. A., AND EBERS, G. C. Onset of secondary progressive phase and long-term evolution of multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry* 85, 1 (2014), 67–75.

- [188] SCALFARI, A., NEUHAUS, A., DEGENHARDT, A., RICE, G. P., MURARO, P. A., DAUMER, M., AND EBERS, G. C. The natural history of multiple sclerosis, a geographically based study 10: relapses and long-term disability. *Brain* 133, 7 (2010), 1914–1929.
- [189] SCANDERBEG, A. C., TOMAIUOLO, F., SABATINI, U., NOCENTINI, U., GRASSO, M. G., AND CALTAGIRONE, C. Demyelinating plaques in relapsing-remitting and secondary-progressive multiple sclerosis: Assessment with diffusion MR imaging. *American Journal of Neuroradiology* 21, 5 (2000), 862–868.
- [190] SCHAEFER, P. W., GRANT, P. E., AND GONZALEZ, R. G. Diffusion-weighted MR imaging of the brain. *Radiology* 217, 2 (2000), 331–345.
- [191] SCHAPIRE, R. E. A brief introduction to boosting. In *Proceedings of the 16th International Joint Conference on Artificial Intelligence - Volume 2* (San Francisco, CA, USA, 1999), IJCAI'99, Morgan Kaufmann Publishers Inc., pp. 1401–1406.
- [192] SCHREIBER, K., SØRENSEN, P. S., KOCH-HENRIKSEN, N., WAGNER, A., BLINKENBERG, M., SVARER, C., AND PETERSEN, H. C. Correlations of brain MRI parameters to disability in multiple sclerosis. *Acta Neurologica Scandinavica* 104, 1 (2001), 24–30.
- [193] SCHUMACHER, G. A., BEEBE, G., KIBLER, R. F., KURLAND, L. T., KURTZKE, J. F., MCDOWELL, F., NAGLER, B., SIBLEY, W. A., TOURTELLOTTE, W. W., AND WILLMON, T. L. Problems of experimental trials of therapy in multiple sclerosis: Report by the panel on the evaluation of experimental trials of therapy in multiple sclerosis. *Annals of the New York Academy of Sciences* 122, 1 (1965), 552–568.
- [194] SCHUMACHER, G. A., BEEBE, G., KIBLER, R. F., KURLAND, L. T., KURTZKE, J. F., MCDOWELL, F., NAGLER, B., SIBLEY, W. A., TOURTELLOTTE, W. W., AND WILLMON, T. L. Problems of experimental trials of therapy in multiple sclerosis: Report by the panel on the evaluation of experimental trials of therapy

- in multiple sclerosis. *Annals of the New York Academy of Sciences* 122, 1 (1965), 552–568.
- [195] SÖDERSTRÖM, M., YA PING, J., HILLERT, J., AND LINK, H. Optic neuritis: prognosis for multiple sclerosis from MRI, CSF, and HLA findings. *Neurology* 50, 3 (1998), 708–14.
- [196] SHARRACK B, H. R. Clinical scales for multiple sclerosis. *J Neurol Sci.* 135, 1 (1996), 1–9.
- [197] SHIM, Y., CHUNG, J., AND CHOI, I. *A comparison study of cluster validity indices using a nonhierarchical clustering algorithm*, vol. 1. 12 2005, pp. 199–203.
- [198] SIGAL, T., SHMUEL, M., MARK, D., GIL, H., AND ANAT, A. Diffusion tensor imaging of corpus callosum integrity in multiple sclerosis: Correlation with disease variables. *Journal of Neuroimaging* 22, 1 (2012), 33–37.
- [199] SIJENS, P. E., IRWAN, R., POTZE, J. H., MOSTERT, J. P., DE KEYSER, J., AND OUDKERK, M. Analysis of the human brain in primary progressive multiple sclerosis with mapping of the spatial distributions using 1H MR spectroscopy and diffusion tensor imaging. *European Radiology* 15, 8 (Aug 2005), 1686–1693.
- [200] SIMON, J. H. Update on multiple sclerosis. *Radiologic Clinics* 44, 1.
- [201] SIMPSON, S., TAYLOR, B., BLIZZARD, L., PONSONBY, A.-L., PITTAS, F., TREMLETT, H., DWYER, T., GIES, P., AND VAN DER MEI, I. Higher 25-hydroxyvitamin D is associated with lower relapse risk in multiple sclerosis. *Annals of Neurology* 68, 2 (2010), 193–203.
- [202] SMOLDERS, J., MENHEERE, P., KESSELS, A., DAMOISEAUX, J., AND HUPPERTS, R. Association of vitamin D metabolite levels with relapse rate and disability in multiple sclerosis. *Multiple Sclerosis Journal* 14, 9 (2008), 1220–1224.

- [203] SOILU-HÄNNINEN, M., AIRAS, L., MONONEN, I., HEIKKILÄ, A., VILJANEN, M., AND HÄNNINEN, A. 25-hydroxyvitamin d levels in serum at the onset of multiple sclerosis. *Multiple Sclerosis Journal* 11, 3 (2005), 266–271.
- [204] SONG, S.-K., SUN, S.-W., RAMSBOTTOM, M. J., CHANG, C., RUSSELL, J., AND CROSS, A. H. Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. *NeuroImage* 17, 3 (2002), 1429 – 1436.
- [205] SORENSEN, A. G., COPEN, W. A., ØSTERGAARD, L., BUONANNO, F. S., GONZALEZ, R. G., RORDORF, G., ROSEN, B. R., SCHWAMM, L. H., WEISSKOFF, R. M., AND KOROSHETZ, W. J. Hyperacute stroke: Simultaneous measurement of relative cerebral blood volume, relative cerebral blood flow, and mean tissue transit time. *Radiology* 210, 2 (1999), 519–527.
- [206] SORMANI, M. P., BONZANO, L., ROCCATAGLIATA, L., CUTTER, G. R., MANCARDI, G. L., AND BRUZZI, P. Magnetic resonance imaging as a potential surrogate for relapses in multiple sclerosis: A meta-analytic approach. *Annals of Neurology* 65, 3 (2009), 268–275.
- [207] SPILKER, J., KONGABLE, G., BARCH, C., BRAIMAH, J., BRATTINA, P., DALEY, S., DONNARUMMA, R., RAPP, K., AND SAILOR, S. Using the NIH stroke scale to assess stroke patients. the NINDS rt-PA stroke study group. *J Neurosci Nurs.* . 29, 6 (1997), 384–92.
- [208] SRINIVASAN, A., GOYAL, M., AZRI, F. A., AND LUM, C. State-of-the-art imaging of acute stroke. *RadioGraphics* 26, suppl_1 (2006), S75–S95.
- [209] STAFF, N., LUCCHINETTI, C., AND KEEGAN, B. Multiple sclerosis with predominant, severe cognitive impairment. *Archives of Neurology* 66, 9 (2009), 1139–1143.
- [210] STEFANO, N. D., MATTHEWS, P. M., AND ARNOLD, D. L. Reversible decreases in N-acetylaspartate after acute brain injury. *Magnetic Resonance in Medicine* 34, 5 (1995), 721–727.

- [211] ØSTERGAARD, L., SORENSEN, A. G., KWONG, K. K., WEISSKOFF, R. M., GYLDENSTED, C., AND ROSEN, B. R. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part II: Experimental comparison and preliminary results. *Magnetic Resonance in Medicine* 36, 5 (1996), 726–736.
- [212] SUAREZ, B., O’ROURKE, D., VAN EERDEWEGH, P., AND OPITZ, J. M. Power of the affected-sib-pair method to detect disease susceptibility loci of small effect: An application to multiple sclerosis. *American Journal of Medical Genetics* 12, 3 (1982), 309–326.
- [213] SWETS, J. ROC analysis applied to the evaluation of medical imaging techniques. *Invest Radiol.* 14, 2 (1979), 109–21.
- [214] SYMON, L. Recovery of brain function after ischaemia. *Acta Neurochir Suppl* 41 (1987), 97–103.
- [215] TASCHLER, B., GE, T., BENDFELDT, K., MÜLLER-LENKE, N., JOHNSON, T., AND NICHOLS, T. Spatial modeling of multiple sclerosis for disease subtype prediction. *Med Image Comput Comput Assist Interv.* 17, 2 (2014), 797–804.
- [216] TEDEHOLM, H., LYCKE, J., SKOOG, B., AND ET AL. Time to secondary progression in patients with multiple sclerosis who were treated with first generation immunomodulating drugs. *Multiple Sclerosis (Houndmills, Basingstoke, England).* 19, 6 (2013), 765–774.
- [217] THAVENDIRANATHAN, P., WALLS, M., GIRI, S., VERHAERT, D., RAJAGOPALAN, S., MOORE, S., SIMONETTI, O. P., AND RAMAN, S. V. Improved detection of myocardial involvement in acute inflammatory cardiomyopathies using T2 mapping clinical perspective. *Circulation: Cardiovascular Imaging* 5, 1 (2012), 102–110.

- [218] THE NATIONAL INSTITUTE OF NEUROLOGICAL DISORDERS AND STROKE RT-PA STROKE STUDY GROUP. Tissue plasminogen activator for acute ischemic stroke. *New England Journal of Medicine* 333, 24 (1995), 1581–1588.
- [219] THE OPTIC NEURITIS STUDY GROUP. Multiple sclerosis risk after optic neuritis: Final optic neuritis treatment trial follow-up. *Archives of Neurology* 65, 6 (2008), 727–732.
- [220] THOMPSON, A., POLMAN, C., MILLER, D., McDONALD, W., BROCHET, B., FILIPPI, M., MONTALBAN, X., AND DE SÁ, J. Primary progressive multiple sclerosis. *Brain* 120, 6 (6 1997), 1085–1096.
- [221] TRABOULSEE, A., SIMON, J., STONE, L., FISHER, E., JONES, D., MALHOTRA, A., NEWSOME, S., OH, J., REICH, D., RICHERT, N., RAMMOHAN, K., KHAN, O., RADUE, E.-W., FORD, C., HALPER, J., AND LI, D. Revised recommendations of the consortium of MS centers task force for a standardized MRI protocol and clinical guidelines for the diagnosis and follow-up of multiple sclerosis. *American Journal of Neuroradiology* 37, 3 (2016), 394–401.
- [222] TRABOULSEE, A. L., AND LI, D. The role of MRI in the diagnosis of multiple sclerosis. *Adv Neurol* 98 (2006), 125–46.
- [223] TRAPP, B. D., PETERSON, J., RANSOHOFF, R. M., RUDICK, R., MÖRK, S., AND BÖ, L. Axonal transection in the lesions of multiple sclerosis. *New England Journal of Medicine* 338, 5 (1998), 278–285.
- [224] TRUYEN, L., VAN WAESBERGHE, J., VAN WALDERVEEN, M., VAN OOSTEN, B., POLMAN, C., HOMMES, O., ADÈR, H., AND BARKHOF, F. Accumulation of hypointense lesions ("black holes") on T1 spin-echo MRI correlates with disease progression in multiple sclerosis. *Neurology* 47, 6.
- [225] VAN DER MEI, I. A. F., PONSONBY, A.-L., DWYER, T., BLIZZARD, L., SIMMONS, R., TAYLOR, B. V., BUTZKUEVEN, H., AND KILPATRICK, T. Past expo-

- sure to sun, skin phenotype, and risk of multiple sclerosis: case-control study. *BMJ* 327, 7410 (2003), 316.
- [226] VASCONCELOS, C. C. F., AND ET AL. Prognostic factors associated with long-term disability and secondary progression in patients with multiple sclerosis. *Multiple Sclerosis and Related Disorders* 8 (2016), 27 – 34.
- [227] VERBEKE, G., AND LESAFFRE, E. A linear mixed-effects model with heterogeneity in the random-effects population. *Journal of the American Statistical Association* 91, 433 (1996), 217–221.
- [228] VILLRINGER, A., ROSEN, B. R., BELLIVEAU, J. W., ACKERMAN, J. L., LAUFER, R. B., BUXTON, R. B., CHAO, Y.-S., WEDEENAND, V. J., AND BRADY, T. J. Dynamic imaging with lanthanide chelates in normal brain: Contrast due to magnetic susceptibility effects. *Magnetic Resonance in Medicine* 6, 2 (1988), 164–174.
- [229] VOLLMER, T. The natural history of relapses in multiple sclerosis. *Journal of the Neurological Sciences* 256 (2007), 5–13.
- [230] VUKUSIC, S., VAN BOCKSTAEL, V., GOSSELIN, S., AND CONFAVREUX, C. Regional variations in the prevalence of multiple sclerosis in French farmers. *Journal of Neurology, Neurosurgery & Psychiatry* 78, 7 (2007), 707–709.
- [231] WAHLGREN, N., AHMED, N., DÁVALOS, A., FORD, G. A., GROND, M., HACKE, W., HENNERICI, M. G., KASTE, M., KUELKENS, S., LARRUE, V., LEES, K. R., ROINE, R. O., SOINNE, L., TONI, D., AND VANHOOREN, G. Thrombolysis with alteplase for acute ischaemic stroke in the safe implementation of thrombolysis in stroke-monitoring study (SITS-MOST): an observational study. *The Lancet* 369, 9558 (2007), 275 – 282.
- [232] WARACH, J. G. M. . S. Imaging of acute stroke. *Nature Reviews Neurology* 6 (2010), 560–571.

- [233] WARLOP, N. P., ACHTEN, E., DEBRUYNE, J., AND VINGERHOETS, G. Diffusion weighted callosal integrity reflects interhemispheric communication efficiency in multiple sclerosis. *Neuropsychologia* 46, 8 (2008), 2258 – 2264.
- [234] WEI, H. C., CHUNG, C. C., AND JEN, L. C. A practical guide to support vector classification. Tech. rep., Department of Computer Science, National Taiwan University, 2003.
- [235] WERRING, D., BRASSAT, D., DROOGAN, A., CLARK, C., SYMMS, M., BARKER, G., MACMANUS, D., THOMPSON, A., AND MILLER, D. The pathogenesis of lesions and normal-appearing white matter changes in multiple sclerosis: a serial diffusion MRI study. *Brain* 123, Pt 8 (2000), 1667–76.
- [236] WERRING, D., CLARK, C., BARKER, G., THOMPSON, A., AND MILLER, D. Diffusion tensor imaging of lesions and normal-appearing white matter in multiple sclerosis. *Neurology* 52, 8 (1999), 1626–32.
- [237] WESTERLIND, H., RAMANUJAM, R., UVEHAG, D., KUJA-HALKOLA, R., BOMAN, M., BOTTAI, M., LICHTENSTEIN, P., AND HILLERT, J. Modest familial risks for multiple sclerosis: a registry-based study of the population of Sweden. *Brain* 137, 3 (2014), 770–778.
- [238] WILLOUGHBY EW, P. D. Scales for rating impairment in multiple sclerosis: a critique. *Neurology* 38, 11 (1988), 1793–8.
- [239] WOLINSKY JS, NARAYANA PA, F. M. Proton magnetic resonance spectroscopy in multiple sclerosis. *Neurology* 40, 11 (1990), 1764–9.
- [240] WU, O., KOROSHETZ, W. J., OSTERGAARD, L., BUONANNO, F. S., COPEN, W. A., GONZALEZ, R. G., RORDORF, G., ROSEN, B. R., SCHWAMM, L. H., WEISSKOFF, R. M., AND SORENSEN, A. G. Predicting tissue outcome in acute human cerebral ischemia using combined diffusion- and perfusion-weighted MR imaging. *Stroke* 32, 4 (2001), 933–942.

- [241] WU, O., ØSTERGAARD, L., WEISSKOFF, R. M., BENNER, T., ROSEN, B. R., AND SORENSEN, A. G. Tracer arrival timing-insensitive technique for estimating flow in MR perfusion-weighted imaging using singular value decomposition with a block-circulant deconvolution matrix. *Magnetic Resonance in Medicine* 50, 1 (2003), 164–174.
- [242] YALDIZLI, Ö., ATEFY, R., GASS, A., STURM, D., GLASSL, S., TETTENBORN, B., AND PUTZKI, N. Corpus callosum index and long-term disability in multiple sclerosis patients. *Journal of Neurology* 257, 8 (Aug 2010), 1256–1264.
- [243] YEO, T. W., DE JAGER, P. L., GREGORY, S. G., BARCELLOS, L. F., WALTON, A., GORIS, A., FENOGLIO, C., BAN, M., TAYLOR, C. J., GOODMAN, R. S., WALSH, E., WOLFISH, C. S., HORTON, R., TRAHERNE, J., BECK, S., TROWSDALE, J., CAILLIER, S. J., IVINSON, A. J., GREEN, T., POBYWAJLO, S., LANDER, E. S., PERICAK-VANCE, M. A., HAINES, J. L., DALY, M. J., OKSENBERG, J. R., HAUSER, S. L., COMPSTON, A., HAFLER, D. A., RIOUX, J. D., AND SAWCER, S. A second major histocompatibility complex susceptibility locus for multiple sclerosis. *Annals of Neurology* 61, 3 (2007), 228–236.

List of publications

Papers in preparation for submission to international journals

1. **Ceren Tozlu**, Brice Ozenne, Tae-Hee Cho, Norbert Nighoghossian, Irene Klærke Mikkelsen, Laurent Derex, Marc Hermier, Salvador Pedraza, Jens Fiehler, Leif Østergaard, Yves Berthezène, Jean-Claude Baron, Delphine Maucort-Boulch. "Comparison of classification methods for tissue outcome after ischemic stroke". Article will be submitted to *Frontiers in Neurology*
2. **Ceren Tozlu**, Dominique Sappey-Marinier, François Cotton, Sandra Vukusic, Françoise Durand-Dubief, Delphine Maucort-Boulch. "Modeling individual disability course of multiple sclerosis patients using clinical and imaging data". Article will be submitted to *Frontiers in Neurology*
3. **Ceren Tozlu**, Françoise Durand-Dubief, François Cotton, Sandra Vukusic, Dominique Sappey-Marinier, Delphine Maucort-Boulch. "Clustering of individual disability progression trajectories of multiple sclerosis patients using clinical and imaging data". Article will be submitted to *Frontiers in Neurology*

Oral presentations at national and international conferences

1. **Ceren Tozlu**, Françoise Durand-Dubief, François Cotton, Sandra Vukusic, Dominique Sappey-Marinier, Delphine Maucort-Boulch. Classification of individual disability progression trajectories of Multiple Sclerosis patients using clinical and multimodal imaging data. Systems Biology at University of Lyon Meeting Days - **Lyon, France**. 13 December 2016
2. **Ceren Tozlu**, Françoise Durand-Dubief, François Cotton, Sandra Vukusic, Dominique Sappey-Marinier, Delphine Maucort-Boulch. Classification of individual disability progression trajectories of Multiple Sclerosis patients using clinical and multimodal imaging data. Groupe de Recherche (GDR) « Statistique et Santé » days organised by *SFdS* (Société Française de Statistique). **Lyon, France**. 27-28 June 2016
3. **Ceren Tozlu**, Brice Ozenne, Tae-Hee Cho, Norbert Nighoghossian, Irene Klærke Mikkelsen, Laurent Derex, Marc Hermier, Salvador Pedraza, Jens Fiehler, Leif Østergaard, Yves Berthezène, Jean-Claude Baron, Delphine Maucort-Boulch. Comparison of classification methods for tissue outcome after ischemic stroke. Applied Statistics International Congress organized by *Statistical Society of Sloveniasur* **Bled, Slovenia**. Sept. 2014

Posters presented at national and international conferences

1. **Ceren Tozlu**, Françoise Durand-Dubief, François Cotton, Sandra Vukusic, Dominique Sappey-Marinier, Delphine Maucort-Boulch. Classification of individual disability progression trajectories of Multiple Sclerosis patients using clinical and multimodal imaging data. Congress of the International Society for Clinical Biostatistics. **Vigo, Spain**. 9-13 July 2017
2. **Ceren Tozlu**, Dominique Sappey-Marinier, François Cotton, Sandra Vukusic, Françoise Durand-Dubief, Delphine Maucort-Boulch. Modeling individual disability course of multiple sclerosis patients using clinical and imaging data. Congress of the International Society for Clinical Biostatistics. **Vigo, Spain**. 9-13 July 2017
3. **Ceren Tozlu**, Françoise Durand-Dubief, François Cotton, Sandra Vukusic, Dominique Sappey-Marinier, Delphine Maucort-Boulch. Classification of individual disability progression trajectories of Multiple Sclerosis patients using clinical and multimodal imaging data. 22nd International Conference on Computational Statistics. **Oviedo, Spain**. 23-26 August 2016

Mrs Ceren TOZLU

Address : 6 rue Mazard, 69002, Lyon, France

Tel : +33 620 66 52 62

Mail : tozluceren@gmail.com

Nationality : Turkish

Date of birth : 05.08.1989



Biostatistician

Education

- Feb.2015 - *Université Claude Bernard & Hospices Civils de Lyon - Lyon, France*
Thesis of science on "Predictive modeling of individual disability progression in Multiple Sclerosis based on longitudinal multimodal Magnetic Resonance Imaging and clinical data". Expected graduation date is March 2018.
- Jan. - June 2014 *Hospices Civils de Lyon - Lyon, France*
Internship in the Department of Biostatistics and Bioinformatics on "Comparison of classification methods (machine learning methods) for tissue outcome after ischemic stroke"
- 2013 - 2014 *Université Claude Bernard - Lyon, France*
Master's degree in Biostatistics, Biomathematics, Bioinformatics and Health (B3S); with scholarship from the French Government
- 2007 - 2013 *Galatasaray University - Istanbul, Turkey*
Bachelor's degree in Mathematics (lessons in French) - Rank 5/32
- 2010 - 2011 *Université de Lille 1 Sciences et Technologie - Lille, France*
ERASMUS exchange programme - Licence 3 of Pure Mathematics: Linear and Affine Algebra, Differential Calculus & Numerical analysis

Teaching activities

- Sep 2015- Dec 2017 • Biostatistics for first-year medical students at École de Santé des Armées – **Bron, France**
- Sep 2015- Dec 2017 • Biostatistics course for master students of Cancer, Neuroscience, Biostatistics and Public Health departments at *Université de Claude Bernard* - **Lyon, France**
- Biostatistics and Modeling, Master 1 level
 - Biostatistics, Master 1 level of Cancer and Master 1 level of Neuroscience
 - Mathematics for Biology and Medicine: Theory and Applications, Master 2 level
 - Generalized linear models and survival models, Master 2 level
- 2008 - 2013 • Part-time work: teaching Mathematics at *Nisantasi High School* - **Istanbul, Turkey**
- July 2012 • Classes at the Science Festival through the implementation of games on mathematics for primary and secondary school students - **Marmaris, Turkey**

Professional experiences

- 17-18 Nov. 2016 • Participation to the statistical workshop organized by *SFdS* (Société Française de Statistique) on "Spatio-temporal processes and trajectory modeling" - **Paris, France**
- 07-08 April 2016 • Participation to the statistical workshop organized by Club Statistiques et Mathématiques Appliquées à la Cancérologie (Cancéropôle Grand Sud-Ouest) on "Biostatistical and mathematical modeling of imaging data" - **Bordeaux, France**
- 05-06 Nov 2015 • Participation to the statistical workshop organized by *SFdS* (Société Française de Statistique) on "Mixed effect models for longitudinal data" - **Paris, France**
- 3 - 14 Sept. 2012 • Participation to the conference organised by *CIMPA* (Centre International de Mathématiques Pures et Appliquées) with Mme Meral TOSUN on "Singularity of spaces, functions, and foliation" - **Fès, Maroc**
- Sept.2011-June2012 • End-of-studies project on "free divisors" under the supervision of Ms Meral TOSUN, Department of Mathematics of Galatasaray University - **Istanbul, Turkey**
- July 2009 • Participation to the summer school organised by *İzmir University of Economics* and *Aziz Nesin Vakfi foundation* on abstract algebra - **Izmir, Turkey**

Publications

- In progress • Article on "Modeling individual disability course of multiple sclerosis patients using clinical and imaging data"
- In progress • Article on "Clustering of individual disability progression trajectories of multiple sclerosis patients using clinical and imaging data"
- In progress • Article on "Comparison of classification methods for tissue outcome after ischemic stroke"

Communications

- 9-13 July 2017 • Two poster presentations at the congress of the International Society for Biostatistics on "Classification of individual disability progression trajectories of Multiple Sclerosis patients using clinical and multimodal imaging data" and "Modeling individual disability course of multiple sclerosis patients using clinical and imaging data" - **Vigo, Spain**
- 13 December 2016 • Oral presentation at Systems Biology at University of Lyon meeting days on "Classification of individual disability progression trajectories of Multiple Sclerosis patients using clinical and multimodal imaging data" - **Lyon, France**
- 27-28 June 2016 • Oral presentation at Groupe de Recherche (GDR) « Statistique et Santé » days organised by *SFdS* (Société Française de Statistique) on "Classification of individual disability progression trajectories of Multiple Sclerosis patients using clinical and multimodal imaging data" - **Lyon, France**
- 23-26 August 2016 • Poster presentation at COMPSTAT 2016 congress (*22nd International Conference on Computational Statistics*) on "Classification of individual disability progression trajectories of Multiple Sclerosis patients using clinical and multimodal imaging data" - **Oviedo, Spain**
- Sept. 2014 • Oral presentation at Applied Statistics International congress organised by *Statistical Society of Sloveniasur* on "Comparison of classification methods for tissue outcome after ischemic stroke" - **Bled, Slovenia**

Awards

- January 2017 & 2018 Open Health Institute - Award for PhD thesis
- September 2014 French Ministry of Foreign and European Affairs - French Ministry of Higher Education and Research Award - Award for a Master's Degree
- 2008-2013 Turkish Education Foundation - Award for a Bachelor's Degree

Skills & Abilities

Languages	Turkish (native), French (fluent), English (good)
Statistical methods	Machine Learning methods (including Support Vector Machines, Artificial Neural Networks, Random Forest, Adaptive Boosting), Predictive modeling of longitudinal data, Linear mixed models, Latent class models, Latent class mixed models, Classification methods on trajectories performed on longitudinal data
Software	Programmation : C, C++, R, and Matlab Tools : Pack Office (excellent), L ^A T _E X(excellent), R (excellent), SAS (excellent), and STATA (excellent)

Professional references

Pr. Delphine Maucort-Boulch (MD - PhD)	Head of the department "Biostatistique et Bioinformatique" of <i>Hospices Civils de Lyon</i> and associated to "Biostatistiques Santé" of "Laboratoire de Biométrie et Biologie Évolutive" at Unité Mixte de Recherche – Centre National de la Recherche Scientifique 5558" at the <i>Université Claude Bernard</i>
(Supervisor of PhD thesis)	e-mail : delphine.maucort-boulch@chu-lyon.fr
Dr. Dominique Sappey-Marinier (PhD)	Researcher and head of the MRI department of the Center for Research in Imaging (CERMEP) and the brain imaging team of the CREATIS research laboratory of CNRS and INSERM at the <i>Université Claude Bernard</i>
(Co-supervisor of PhD thesis)	e-mail : dominique.sappey-marinier@univ-lyon1.fr
Pr. Pascal Roy (MD - PhD)	Head of research team "Biostatistiques Santé" of "Laboratoire de Biométrie et Biologie Évolutive" at Unité Mixte de Recherche – Centre National de la Recherche Scientifique 5558" at the <i>Université Claude Bernard</i>
	e-mail : pascal.roy@chu-lyon.fr
Dr. Claire Bardel-Danjean (PhD)	Teacher of statistics at Université Claude Bernard and Researcher & Hospital Practitioner at <i>Hospices Civils de Lyon</i>
	e-mail : claire.bardel-danjean@univ-lyon1.fr