# Diffusion Magnetic Resonance as the Basis of Novel Biomarkers Using Multidimensional Statistics in the Brain

PhD thesis summary

**Gyebnár Gyula**

János Szentágothai Doctoral School
Semmelweis University

Supervisor:

Lajos Rudolf Kozák MD, PhD

Opponents:

András Jakab MD, PhD

Pál Kaposi Novák, MD, PhD

Head of the Final Examination Committee:

Kinga Karlinger MD, PhD

Members of the Final Examination Committee:

Miklós Krepuska, MD, PhD

Regina Deák-Meszlényi, PhD

Budapest

2020

# I. INTRODUCTION

The last two decades' developments of magnetic resonance imaging (MRI) have propelled the investigation of the central nervous system (CNS) at an unprecedented rate. With the increasing availability of high–quality equipment and the inventions of novel imaging and processing techniques, MRI has advanced to be in the forefront of brain research.

The ever–growing amount of multimodal imaging data, especially with the introduction of diffusion and functional MRI, has been driving the development and use of novel processing and statistical methods. Since different modalities, even distinct MR contrast mechanisms, provide complementary information on brain structure and function, methods combining data from various approaches could aim for more accuracy and sensitivity by leveraging their advantages.

The work described in the Thesis was aimed at developing statistical methods and new biomarkers using combinations of state–of–the–art neuroimaging techniques with diffusion MRI in two separate research projects: one investigating the potential role of diffusion tensor imaging (DTI) in the differential diagnosis of mild cognitive impairment (MCI), and one introducing a novel approach for single subject evaluation that searches for lesion voxels as outliers – tested on data of epilepsy patients with malformations of cortical development.

## 1. Diffusion Magnetic Resonance Imaging

Diffusion magnetic resonance imaging (dMRI) is a general term referring to a group of methods widely used for the non–invasive examination of biological samples and porous materials. By using spatially varying magnetic fields (referred to as diffusion encoding gradients), the measured MR signal is made sensitive to the random thermal motion of water molecules.

Even though diffusion only results in displacements on the microscopic scale, the degree at which water molecules can move in the extracellular environment may be hindered, resulting in restricted, and in many cases anisotropic diffusion. Therefore, the measured signal attenuation in biological samples reflects properties of tissue microstructure.

DTI is the most widely known (and simplest) technique that is capable of handling anisotropy information. DTI is used in both clinical neuroradiology and for innumerable research questions about the CNS, and it was also the method of choice in both research projects of the present Thesis.

In the DTI representation, the 3D Gaussian model uses a 3–by–3 tensor $\boldsymbol{D}$ to describe diffusion anisotropy, which can be viewed as an ellipsoid. The lengths of the ellipsoid's half–axes are the eigenvalues ($\lambda_1$, $\lambda_2$, $\lambda_3$) of $\boldsymbol{D}$ and the corresponding eigenvectors ($\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$) determine it's orientation. DTI can be utilized with two distinct approaches: the orientation information facilitates tractography, while the eigenvalues can be used to derive measures of anisotropy and diffusivity that reflect properties of tissue microstructure. The most well–known are mean diffusivity (MD) that reflects the volume of the extracellular space, and fractional anisotropy (FA), generally viewed as a measure of fiber coherence in the white matter (WM).

## 2. Statistics in brain MRI studies

Numerous approaches has been employed to utilize the superior image quality and various contrast mechanisms of MRI for studying brain structure, morphology, maturation, and the effects of various diseases. The simplest, univariate methods measure the size or volume of specific structures, e.g. the hippocampus and either compare it between individuals or cohorts of different diseases with healthy control subjects; or correlate it with specific, often neuropsychology–related measures.

Explorative methods also emerged for identifying systemic effects, using either structural or functional atlases, or common coordinate systems, called templates, that facilitate multi-subject statistical inference on the voxel level. Such methods are called voxel–based analyses.

Performing a large number of statistical tests simultaneously (mass univariate testing) inherently degrades the reliability of inference – leading to an issue referenced as the problem of multiple comparisons. Remedies can be: (1) reducing the number of the tests by the limiting the number of examined structures or volumes, (2) utilizing the fact that the values measured in the brain are not independent, or (3) applying conservative thresholds for inference in order to control the rate of false positives.

Multidimensional and multimodal studies aim to combine information from independent sources in order to raise statistical power; a feat sought after in the neuroimaging literature. Several strategies were employed to implement such combination at different levels of statistical analysis throughout the past two decades, using (and sometimes combining) voxel–wise, surface–based, or region–level methods.

## 3. The Mahalanobis–distance

The Mahalanobis–distance is a measure of dissimilarity, commonly used in multivariate outlier detection problems, that one may view as a multidimensional one–sample *T*–statistic. We employed the squared Mahalanobis–distance ($D^2$) in our second study as the basis of lesion detection, searching for abnormal voxels as outliers when comparing a single subject to a group of controls in the three dimensional space of DTI eigenvalues.

## 4. Research topics – Clinical importance

### *4.1. Mild Cognitive Impairment*

Alzheimer's Disease (AD) is the most common neurodegenerative disorder among the aging population, which has a substantial social and economic burden in western societies such as countries of the European Union or the United States.

While there is no effective treatment for AD at the moment, future interventions may likely be effective in an early stage of the disease. Therefore, the intermediate stage between the mild decrease of cognitive functioning in physiological aging and the severe decline in dementia known as 'mild cognitive impairment' has gained a lot of interest in the last decade. Subtypes of MCI can be differentiated such as the amnestic (aMCI – with high conversion rate to AD), and the non–amnestic (naMCI – associated with other dementia variants) subtypes. The conversion rate from aMCI to AD is much higher, which underlines the significance of differentiation between these subtypes.

### *4.2. Drug resistant epilepsies (DREs)*

Drug resistance affects about *20–30%* of the epileptic patient population, causing severely impaired quality of life and a difficult to treat situation. Most of the drug resistant cases (*~60%*) are focal epilepsies; nevertheless, there are generalized forms. Malformations of cortical development (MCDs) are among the most frequent etiological factors causing DRE. Patients are often candidates for surgical intervention; however, the probability of postoperative seizure freedom is remarkably lower in cases lacking any identifiable lesions on conventional MRI; better visualization can be crucial for improving surgical outcomes.

# II. AIMS

## 1. DTI and Mild Cognitive Impairment

The primary aim of the first study was to find the possible differences between the subgroups of MCI, which may increase prognostic capability at an early stage, and a further aim was to confirm the recent findings regarding the DTI differences observed between controls and MCI subjects. Based on previous evidence the most prominent between group differences and the strongest correlations with memory functions were expected in the cingulum and the fornix. With the available literature regarding DTI findings in MCI as reference, we also aimed to demonstrate the feasibility of the whole brain voxel-level analysis using a study-specific template, retaining statistical power even with the conservative correction for multiple comparisons.

The secondary aim of the study was to determine in which brain regions can DTI measurements provide added benefit when performed in combination with volumetric examinations, to help the differentiation between patients with MCI and healthy subjects.

## 2. Mahalanobis–distance in MCD lesion detection

The aim of the second study was to evaluate the performance of a novel, Mahalanobis–distance–based statistical approach using DTI data, – for detecting microstructural abnormalities by simulations using data from standard multivariate normal distribution and from healthy controls. Based on the simulation results we also aimed to demonstrate the potential clinical utility of the approach in select cases of patients with MCDs.

# III. METHODS

## 1. DTI and Mild Cognitive Impairment

Briefly, in this study different statistical approaches were applied to (a) identify those white matter structures which are the most sensitive to early impairment in pathological aging, and (b) to estimate if diffusion metrics can extend the differentiation performance of volumetry.

The dMRI and $T_1$–weighted imaging data of *65* subjects (*18* with amnestic MCI, *20* with non–amnestic MCI, and *27* healthy controls) acquired at *3T* was analyzed retrospectively, along with the results of neuropsychological evaluations (Rey Auditory Verbal Learning Test – RAVLT, Addenbrooke's Cognitive Examination – ACE, Mini Mental State Examination – MMSE, and Paired Associates Learning – PAL). Subjects included in the study were categorized as aMCI, naMCI, and healthy controls according to the Petersen criteria.

Data processing and statistical tests were performed using specific, Matlab-based toolboxes (ExploreDTI and SPM12), in-house algorithms designed for the study, and SAS. Maps of FA and MD were transformed into a common, study–specific coordinate system using the 'DARTEL' tools of SPM12.

We performed voxelwise correlation analyses between neuropsychological tests and the DTI parameters to assess whether these tests capture the examined aspects of cognitive performance and that the DTI metrics reflect the state of tissue microstructure in relation to them. Next, we compared groups of healthy individuals, and at–risk subgroups of amnestic and non–amnestic MCI on the voxel–level to solidify the results of the correlation analyses and to identify the regions showing significant between–group differences.

We performed both correlation and between–group analyses in predefined regions of interest (*36* ROIs of the 'JHU White–Matter Atlas') as well. With less independent tests to perform and thereby using more liberal thresholds for multiple comparisons correction, the ROI–level approach may exhibit higher sensitivity. Even more so as DTI is most sensitive in the WM, ROI–level results may prove to be more stable, rendering this approach better suited for discriminative models.

Finally, in order to prove our hypothesis that DTI measures of these regions can improve the differentiation performance achievable with grey matter (GM) volumetry, stepwise logistic regression analysis was performed with a K–fold cross–validation approach.

In the first set of models, only cortical thickness measurements and subcortical brain structure volumes (namely the volume of the hippocampus, the cortical thickness of the entorhinal cortex, the fusiform gyrus, the precuneus, and the isthmus of the cingulate gyrus) were entered as variables. These volumetric and thickness measurements of GM structures were amended with the FA and MD measurements of *36* WM ROIs (i.e. tracts) in the second and third sets of models, separately.

## 2. Mahalanobis–distance in MCD lesion detection

Diffusion and $T_1$–weighted MR imaging data of *45* healthy control subjects and *13* patients with MCDs was acquired at 3T, and used retrospectively, along with 2D fluid attenuated inversion recovery (FLAIR) images for aiding visualization. MCD subtypes included polymicrogyria (in two patients) schizencephaly (two patients), subependymal heterotopia (in three patients), focal cortical dysplasia (in six patients), cortical dysgenesis (in three patients) and other, not clearly identifiable malformations (in four patients). Several other types of abnormalities were also identified in the patient group, such as DNT (in one patient, later confirmed by histopathology), ischemic WM lesions (in two patients), a gliotic cyst (in one patient), focal gliosis (in one patient, also confirmed by subsequent histopathology), hippocampal sclerosis (in four patients) and malrotation of the hippocampus (in one patient). Diagnoses of MCD subtypes were based on neuroradiology report.

The dMRI data was processed with ExploreDTI, and we used the DARTEL method once again for the group–level coregistration of the eigenvalue images, but the template was created from the $T_1$–weighted images of only the control subjects; patient data was subsequently registered to this common space. The resulting coregistered whole brain DTI eigenvalue images of the healthy subjects were used for three purposes: (a) as data basis for simulations in a ‘bootstrap’ manner, (b) in a leave–one–out examination to measure the performance of coregistration and its effect on false positives, and (c) as reference when patient data was examined.

As part of our epilepsy post–processing protocol we also used the MAP07 toolbox that performs single subject vs. control group comparisons on volumetric $T_1$ data. After re-evaluation, these results, along with additional ROIs created manually to cover lesions that were not identified, served as reference in further analysis.

We have implemented the calculation of the voxel–wise Mahalanobis–distance from the DTI eigenvalue maps, the statistical inference based on

analytically–derived critical values (with false discovery rate – FDR – or family-wise error rate – FWE – correction to address the problem of multiple comparisons), and cluster size thresholding in Matlab scripts and functions.

The performance of the lesion detection method was evaluated using simulations with two distinct sets of data: (a) Gaussian random images and (b) real diffusion tensor eigenvalue maps. Alternative free–response receiver–operator characteristics (AFROC) analyses were carried out on both sets. Simulations were carried out with different contrast–to–noise ratios (CNR) and lesion sizes, with a variable cluster size threshold for controlling the rate of false positives.

The demonstrated high sensitivity makes the approach susceptible to registration artefacts and strong individual variability, resulting in false positives. To measure the impact on patient evaluation, data of the control subjects was used in a leave–one–out examination, comparing each individual to the remaining *44*.

In order to distinguish false positives and increase specificity, clusters were subjected to additional post–processing. From each individual's Tissue Probability Maps (resulting from the DARTEL–pipeline), we defined a new parameter ($\delta$), describing voxel position in the [*-1*, *1*] range, with positive values signaling voxels closer or belonging to the cerebrospinal fluid (CSF), and negative values indicating voxels closer or belonging to WM. Based on the results of leave–one–out examination of controls, a cutoff value was determined, signaling clusters with the majority of voxels from the CSF.

Representative cases of MCDs and other abnormalities were analyzed in the same manner as described above. Results were qualitatively evaluated by comparing the anatomical images and $D^2$ 'heatmaps' along with the *Z*–scored junction maps of MAP07. Clusters of outlying diffusion profile, remaining after the thresholding and artefact removal steps were considered true positive, when good spatial concurrence with the underlying pathology (as observed on anatomical scans) and the reviewed and corrected results of the MAP07 toolbox was ascertained. An additional step included the calculation of the clusters' centers of mass (using the $D^2$–values as weights), and their (physical) distance from the predefined lesion masks.

# IV. RESULTS

## 1. DTI and Mild Cognitive Impairment

### *1.1. Voxelwise analyses*

Voxelwise correlation of the FA and MD values with the results of the four neuropsychological tests was found to be significant ($p < 8.4 \times 10^{-7}$) in several clusters with FDR and FWE correction. The most significant correlation between FA values and the PAL test results was identified in the pars triangularis of the right inferior frontal gyrus. Correlation of MD with three of the four tests was found to be significant in several clusters. The RAVLT score was found to be correlated to MD in the left parahippocampal gyrus, the ACE score correlated significantly to MD in the left parahippocampal gyrus and in the pole of the left middle temporal gyrus. Ten clusters with significant correlation between the Trail Making test and the MD values were identified: one in the angular gyrus and one in the right superior temporal gyrus showed the strongest correlation. The small clusters of significant correlations ($p$–values in the order of $10^{-7}$) were identified as peaks of trend–like behavior.

Between-group comparison of mean diffusivity values yielded several significant results with FWE-correction, while no regions showed significantly different FA. One–way ANOVA confirmed MD differences in seven clusters of voxels, with p–values below $2.3 \times 10^{-6}$. Seventeen regions showed increased MD in patients with aMCI compared to controls, a left inferior temporal cluster and a right middle temporal cluster were the largest and most significant. Higher MD in the naMCI group compared to the controls was confirmed with post–hoc T–tests, in a precuneal and a smaller temporal cluster ($p < 8.5 \times 10^{-9}$). No significant difference was identified between the DTI scalar values of the two groups of MCI patients on the voxel level.

As expected, all the regions deemed significant with FWE correction emerged as peaks of the $T$–score "landscape": focal points of regions with trend–like MD differences.

### *1.2. ROI–based analyses*

Both FA and MD in the left cingulum and in the left stria terminalis / left crus of the fornix correlated with the RAVLT total scores, with the total adjusted trials in the PAL test, and with the total and verbal fluency scores of the ACE. After correction for age the correlation between RAVLT total score and MD in the left cingulum ($p = 0.0008$), and the correlation between the PAL test result and MD in the left stria terminalis / left crus of the fornix ($p = 0.0001$) remained significant. Furthermore, the correlations of the RAVLT ($p = 0.0004$) and the verbal fluency subscore of the ACE ($p = 0.001$) with FA in the left cingulum were found significant.

FA of the left cingulum (hippocampal subdivision) was significantly lower in the aMCI group relative to the control group ($p < 0.0001$) and to the naMCI group ($p < 0.0004$). MD of the left cingulum was significantly higher in the aMCI group relative to controls ($p < 0.0001$) and subjects with naMCI ($p < 0.0012$). Furthermore, a tendency level difference in MD was detected between controls and aMCI patients in the left stria terminalis and the left crus of the fornix, while neither FA nor MD was found to be significantly different between controls and naMCI subjects ($p > 0.05$).

### *1.3. ROI–based logistic regression analysis for differentiation between study groups*

For the differentiation between aMCI subjects and healthy controls, the volume of the left or right hippocampus (extended by the cortical thickness of the precuneus in three subsets) resulted in *78.95%* correct categorization, which could not be improved significantly with the inclusion of any of the DTI-based measures.

Differentiation between naMCI subjects and healthy controls found the average cortical thickness of the precuneus and the volume of the right hippocampus meaningful when working with the volumetry- and thickness-measures, but with only *52.63%* correct decisions. Adding the FA measures to the models resulted in two more correct decisions (*55.26%)*, while MD measures yielded three additional correct decisions (*57.89%*) in total, with no region evidently marked to be meaningful.

When differentiating between the aMCI and naMCI subjects, in the first set of models, the volume of the left hippocampus was the only meaningful effect, achieving an overall *63.89%* correct categorization performance.

After adding the average FA value of the ‘stria terminalis / left crus of the fornix’ correct categorization with all but one test subsets was

increased, resulting in an overall *86.11%* categorization performance, meaning *22.22%* increase compared to the solely volumetry–based models.

Average MD values of the body of the corpus callosum or the column and body of the fornix also improved the categorization performance in the final sets of models: together with the volume of the left hippocampus, *75%* of subjects were categorized correctly.

## 2. Mahalanobis–distance in MCD lesion detection

### *2.1. Simulations with Gaussian data*

No false positive clusters were identified with size thresholds larger than *4* voxels, meaning that for simulated lesions with voxel values from standard Gaussian distribution, and sizes that are reasonable to assume any true malformation would have, the method had *100%* specificity. Area under the curve (AUC) values of the ROC-plots exceeded *84%* in lesion detection, with lesion sizes above *19* voxels and CNR above *1* FWHM, and more than half of the lesion voxels were identified with $CNR > 3\sigma$, with all lesion sizes and critical values.

### *2.2. Real Eigenvalue simulations*

The false positive rate was decreased to *0.1 – 0.3%* with cluster size thresholds of *6* (with FWE–correction) or *7* (with FDR–correction) voxels. Lesion identification performance was above 70% with *CNR = 1FWHM* in cases of lesions larger than *50* voxels, or with $CNR = 3\sigma$ and at least *35* voxels, using either FDR, or FWE–corrected critical values. More than half of the lesion voxels were identified at *CNR = 2FWHM*, achieving *77.3 – 85.9%* AUC with FDR–corrected, and *75.1 – 77.8%* AUC with FWE corrected critical values. Based on these results, a *7* voxel cluster size threshold was determined for the examinations of real subjects.

### *2.3. Leave–one–out analysis of controls*

The FDR–corrected critical values resulted in an average of *21.11 (5 – 55)* clusters/subject, while the more conservative FWE–correction yielded *4.93 (0 – 13)* clusters in average, *1.79 (0 – 5)* of those being in the WM. After removing clusters based on the $\delta$–values, the number of remaining clusters decreased to an average of *2.79 (0–7)* with an average size of *16.21* voxels *(7–167)*, meaning, that most of those resulting from insufficient registration or normal differences in gyrification patterns (mainly located in the CSF) were filtered out.

### *2.4. Patient Examination*

After applying the Mahalanobis–distance–based lesion detection method with the above–described processing steps to the *16* examinations of the *13* patients, on average *59.4 (35 – 90)* clusters per subject were identified with an average size of *31.4 (7 – 680)* voxels. The majority of these clusters were obvious artefacts, identifiable by their shape and location (e.g. in the occipital lobes, close to and following the GM–CSF boundary, independent of the underlying gyral and sulcal pattern).

Regions with outlying diffusion properties, corresponding to *22* (out of the *23*) MCDs and other abnormalities were identified, in good spatial concurrence with the neuroradiological evaluation and the previously defined lesion masks: the average (physical) distance for clusters deemed positive was *12.07 mm*, in agreement with results from literature.

# V. CONCLUSIONS

## 1. DTI and Mild Cognitive Impairment

The findings of the first study supported the hypothesis that impairments of white matter integrity appear as early signs of pathological cognitive decline in both amnestic and non–amnestic clinical manifestations of MCI. DTI measurements in the fornix, the stria terminalis and the cingulum can add valuable information to analyses based on grey matter volumetry and thus can help detect Alzheimer Disease in an early, preclinical stage.

Future medications in AD are expected to be effective in such early stages, therefore extending volumetric analyses with DTI–derived metrics may help the early identification of the disease well within the possible therapeutic window of these proposed medications to help preventing further decline.

Aside from the prospective advantages of the proposed methods, differentiation between MCI patient groups was also demonstrated to be improved after extending GM volumetry–based models with DTI measurements, yielding an immediate utility of the addition of DTI–based metrics in the evaluation of MCI.

## 2. Mahalanobis–distance in MCD lesion detection

The Mahalanobis–distance based method, proposed in the second study, efficiently combined information from maps of the three diffusion tensor eigenvalues on the voxel–level. Altered diffusion profiles corresponding to malformations of cortical development in single subject vs. control group examinations were detected as outlier values in the voxelwise multidimensional distributions, based on, but not necessarily limited to DTI data.

Searching for pathological brain regions of individuals as outliers, using the Mahalanobis–distance in evaluation of diffusion weighted imaging data (even with more sophisticated models for processing, if necessary) seems to be a viable approach, and as the calculations could easily cover data from other modalities, this evaluation method may substantially advance the field of quantitative MRI in general.

## 3. General conclusions

Multidimensional and multi–modal approaches in the processing and evaluation of brain MR images proved to be efficient in detecting the disease–related alterations of tissue microstructure based on DTI data, both when characterizing impairments associated with cognitive decline and when searching for malformations related to drug resistant epilepsies.

The application of state–of–the–art data processing methods, thorough image correction and using anatomical scans as targets for registration and "up–sampling", combined with the volume–based creation of study–specific templates using the DARTEL method (originally developed for volumetry) was efficient in treating voxel–level dMRI data.

Resulting sample distributions had low observed variance and good spatial congruency between control subjects, which facilitated strong inferences, even when considering the high number of simultaneously performed statistical tests. Voxel–level correlation analyses and between–group comparisons pinpointed the key structures affected through cognitive impairment; these findings subserved the application of multi–modal logistic regression on the region–level and the multidimensional statistics in single subject evaluation.

The combined information resulted in high sensitivity, yielding improved accuracy in the detection of abnormal tissue microstructure, originating from MCI–related changes or epilepsy–related malformations, suggesting the clinical utility of such evaluation methods.

# THE CANDIDATE'S BIBLIOGRAPHY

## Publications related to the thesis

*Gyebnar G, Szabo A, Siraly E, Fodor Z, Sakovics A, Salacz P, Hidasi Z, Csibri E, Rudas G, Kozak LR, Csukly G (2018) What can DTI tell about early cognitive impairment? - Differentiation between MCI subtypes and healthy controls by diffusion tensor imaging. Psychiatry Res Neuroimaging, 272: 46-57.*
*https://doi.org/10.1016/j.pscychresns.2017.10.007*
*IF: 2.270*

*Gyebnar G, Klimaj Z, Entz L, Fabo D, Rudas G, Barsi P, Kozak LR (2019) Personalized microstructural evaluation using a Mahalanobis-distance based outlier detection strategy on epilepsy patients' DTI data - Theory, simulations and example cases. PLoS One: 14(9):e0222720.*
*https://doi.org/10.1371/journal.pone.0222720*
*IF: 2.740*

## Publications unrelated to the thesis

*Lakatos A, Kolossvary M, Szabo M, Jermendy A, Barta H, Gyebnar G, Rudas G, Kozak LR (2019) Neurodevelopmental effect of intracranial hemorrhage observed in hypoxic ischemic brain injury in hypothermia-treated asphyxiated neonates - an MRI study. BMC Pediatr,19(1):430.*
*https://doi.org/10.1186/s12887-019-1777-z*
*IF: 1.909*

*Csukly G, Szabó Á, Polgár P, Farkas K, Gyebnár G, Kozák LR, Stefanics G (2020). Fronto-thalamic structural and effective connectivity and delusions in schizophrenia: A combined DTI/DCM study. Psychological Medicine, 1-11.*
*https://doi.org/10.1017/S0033291720000859*
*IF: 5.813 (2019)*