UCLouvain

École polytechnique de Louvain

Metrics Distribution Analyses - Definition and evaluation of a new method for the microstructural characterisation of the brain of stroke patients.

Author: **Louis Lovat**
Supervisors: **Benoît Macq, Laurence Dricot**
Readers: **Nicolas Delinte, Yves Vandermeeren**
Academic year 2022–2023
Master [120] in Mechanical Engineering

Abstract

Stroke is one of the world's leading causes of disability. The use of magnetic resonance imaging (MRI), and more specifically of diffusion magnetic resonance imaging (dMRI), made it possible to carry out an in-depth analysis of the impact of stroke on the brain of victims. The development of diffusion models such as DTI, NODDI, DIAMOND and MF has enabled precise in-vivo characterisation of brain microstructures. With it, we have acquired the ability to link the changes in these microstructures to their visible effect in patients motor skills using clinical tests. There are many ways of analysing the results of diffusion models, most of which coarsely average the values of metrics over entire brain regions. The aim of this study is to establish a new method called *Metrics Distribution Analyses* (MDA), which aims to better represent diffusion metrics in a region of interest. We tested this method using a database composed of 104 observations from 32 patients suffering from ischaemic stroke, each observation consisting of a diffusion image, a T1-weighted image and a result from a clinical test: the *6-minute walk test*. Using these data, we computed diffusion metric maps over the corticospinal tract of patients and demonstrated the effectiveness of our analyses. We also compared the new method with more traditional ones to put the results in perspective. Our results showed great potential in the use of MDA with the DTI and DIAMOND metrics and less interest in the NODDI and MF metrics. Even though the statistical significance of some of our results suffers from the limitations of this study and could be improved, we concluded that the use of MDA together with other traditional methods could undoubtedly improve the analysis of microstructures based on diffusion images.

Keywords : diffusion MRI (dMRI), corticospinal tract (CST), motor stroke, ischaemic stroke, Metrics Distribution Analyses (MDA)

Acknowledgement

I would like to thank my supervisors Pr Benoît Macq and Pr Laurence Dricot for their patience, trust and support throughout this thesis.

I would also like to thank Nicolas Delinte for his constant availability, his advice and his help.

For allowing me to work on their data, I would also particularly like to thank Pr Yves Vandermeeren and his team, without whom this work would not be possible.

Finally, I would like to thank my family and friends for standing by my side in difficult times, for their trust, their support and for the interest they have shown in my work.

List of Abbreviations

AD	Axial Diffusivity
ADC	Apparent Diffusion Coefficient
aM	Metric asymmetry
ARAT	Action Reach Arm Test
CCD	Cumulative Curve Difference
CSF	Cerebrospinal fluid
CST	Corticospinal Tract
CT	Computed Tomography
dMRI	diffusion Magnetic Resonance Imaging
DIAMOND	DIstribution of Anisotropic MicrOstructural eNvironments in Diffusion
DTI	Diffusion Tensor Imaging
DTW	Dynamic Time Warping (distance)
DW-MRI	Diffusion-Weighted Magnetic Resonance Imaging
FA	Fractional Anisotropy
FID	Free Induction Decay
FOV	Field Of View
KNN	K-Nearest Neighbours
MCD	Mean Curve Distance
MD	Mean Diffusivity
MDA	Metrics Distribution Analyses
MF	Microstructure Fingerprinting
MI	Mutual Information
MLP	Multi Layer Perceptron
MRI	Magnetic Resonance Imaging
6MWT	6-Minute Walk Test
NMR	Nuclear Magnetic Resonance

NODDI	Neurite Orientation Dispersion and Density Imaging
ODI	Orientation Dispersion Index
PET	Positron Emission Tomography
RD	Radial Diffusivity
RF	Radio-Frequency
rM	Metric ratio
ROI	Region Of Interest
SCD	Standard deviation of the Curve Distance
SIAS	Stroke Impairment Assessment Set

List of Figures

Contents

Introduction

Stroke is one of the top leading causes of death in the world, killing every year around 6 million people worldwide. But stroke does not just kill people. Survivors are often left severely disabled and up to one in two surviving victim remains chronically disabled [1, 2, 3]. In order to provide better care for patients affected by this pathology, the study of the impact of strokes and rehabilitation treatments on brain microstructures has developed considerably and now represents a major area of research.

Involved in this research field, new imaging techniques have been developed. These include diffusion magnetic resonance imaging (dMRI). This imaging technique shows the movement of water molecules in tissues caused by diffusion. Based on this principle, mathematical models have been developed that attempt to characterise the microstructure of brain tissues on the basis of diffusion images. The best known of these is the Diffusion Tensor Imaging (DTI) [4, 5, 6].

The use of these models is now commonplace in the analysis of brain microstructures after a stroke. In concrete terms, these models provide metric maps, which consist of assigning a value to each voxel in a diffusion image. These metrics can provide different information about the underlying microstructure, depending on the model used. It is then necessary to synthesise the information from these maps on particular brain regions to derive usable and interpretable data. The Metrics Distribution Analyses (MDA) are a new type of analysis that attempts to better represent the values of a metric distributed over a region of interest, in comparison with methods currently used.

In this study, we define and investigate the new method of the MDA. In the first sections, we begin by re-establishing the theoretical framework necessary for this study by reviewing the literature about magnetic resonance imaging and strokes. We then explain how this study was carried out and how we validated the newly introduced MDA method based on a sample of 104 dMRI images of patients suffering from stroke. To conclude, an analysis and discussion of our results are proposed, including the conclusions and limitations of this work as well as an outlook towards future improvements.

1 State of the art

1.1 Stroke pathology

As it is the centre of interest of this study, it is necessary to truly dive into the understanding of stroke pathology: its extent, its influence on people's health, the mechanisms involved and its materialization in people's brain's microstructure.

1.1.1 Epidemiology

Every year, stroke kills around 6 million people worldwide. It is one of the top leading causes of death in the world. On top of its mortality, stroke also leads to serious disabilities. As a result, up to 50% of stroke survivors remain chronically disabled [1, 2, 3].

There are two main types of stroke: ischaemic and haemorrhagic. The first category is the one being studied here and is caused by a lack of oxygen and glucose in the tissues. It accounts for about 80% of strokes and is often caused by embolism or thrombosis. Yet, up to 40% of ischaemic strokes remain classified as cryptogenic, which refers to unknown underlying causes. The second category, which accounts for the remaining 20% of strokes, is caused by bleeding, provoking an accumulation of blood within the cranial vault [1, 2, 3, 7].

1.1.2 Risk factors

There are many factors that may have an impact on the occurrence of stroke. Some of those factors are partially controllable and account for 82% of stroke causes according to the National Heart, Lung and Blood Institute in the USA [8].

Among the so-called "controllable" risk factors, we find [9]:

`Hypertension:` It is the most important modifiable risk for ischemic stroke with a relative risk reaching 4 in case of severe hypertension (systolic blood pressure above 160 mmHg and/or diastolic blood pressure superior to 95 mmHg).

`Lifestyle:` It combines diet, obesity, alcohol consumption and physical activity. These factors have both indirect impacts on the occurrence of stroke, through an augmentation of the blood pressure for example, and direct impacts, for instance, through fat deposits in vessels related to a bad diet and obesity.

`Cardiac Diseases:` Various cardiac diseases have been proven to have an impact on the occurrence of strokes, some of them being treatable. The most recognized precursor of stroke is atrial fibrillation with an estimate of 50% of directly related cardioembolic strokes. While not being exhaustive, we can also mention cardiac valve abnormalities, mitral annular calcification, myocardial disease and many more as risk factors.

`Diabetes:` On top of related hypertension, obesity and so on, people with diabetes have been confirmed to be prone to a 1.8 to 3 relative risk for an ischaemic stroke.

`Cigarette Smoking:` Smoking cigarette tends to multiply the ischaemic stroke risk by nearly two.

Aside from those controllable risks, other non-modifiable factors may influence the incidence of stroke. These include [8, 9]:

Age: Age is the most important risk factor, doubling the risks of stroke for each successive 10 years after age 55.

Gender: Men are 1.25 times more likely to have a stroke than women.

Ethnicity: Even if this factor may be related to underlying differences between groups (quality of life or other environmental factors), belonging to an ethnic group may influence the stroke incidence by a factor 2.

1.1.3 Pathophysiology

As mentioned above, strokes can be divided into two categories depending on the cause and effect of the stroke. We have chosen to look more closely at the mechanisms involved in ischaemic stroke, as this is the type of stroke which constitutes the data used for this study.

Ischaemic stroke is characterised by blood hypoperfusion. The reduction of oxygen and glucose supply below a critical threshold leads to loss of cellular function and rapid cell death. Within a few minutes, the lack of oxygen and glucose causes mitochondrial and ion pump dysfunction, resulting in cellular plasmic membrane breakdown, swelling of organelles and cellular content into the extracellular medium, and the appearance of a necrotic area called the ischaemic core. In this area, the damage is irreparable, with permanent loss of the affected neurons and other cells. After a stroke, the ischaemic core is mainly exclusively constituted of cerebrospinal fluid (CSF).

The chemical reactions induced by the stroke also cause inflammation in and around the ischaemic core. The area affected by the inflammation is usually called the "penumbra", as seen in Figure 1.1. The number of immune cells in the penumbra increases dramatically. The area is characterised by the presence of damaged but relatively viable neurons and glial cells. Cells in the penumbra undergo reversible death processes such as apoptosis and autophagy. These processes can be halted if oxygen and glucose supply are properly restored and inflammation properly managed. Partial or total recovery of neuronal function is then possible. Until then, the permeability of the blood-brain barrier in the area is increased, sometimes reaching a breaking point, leading to oedema and to further complications [1, 2, 3, 10].

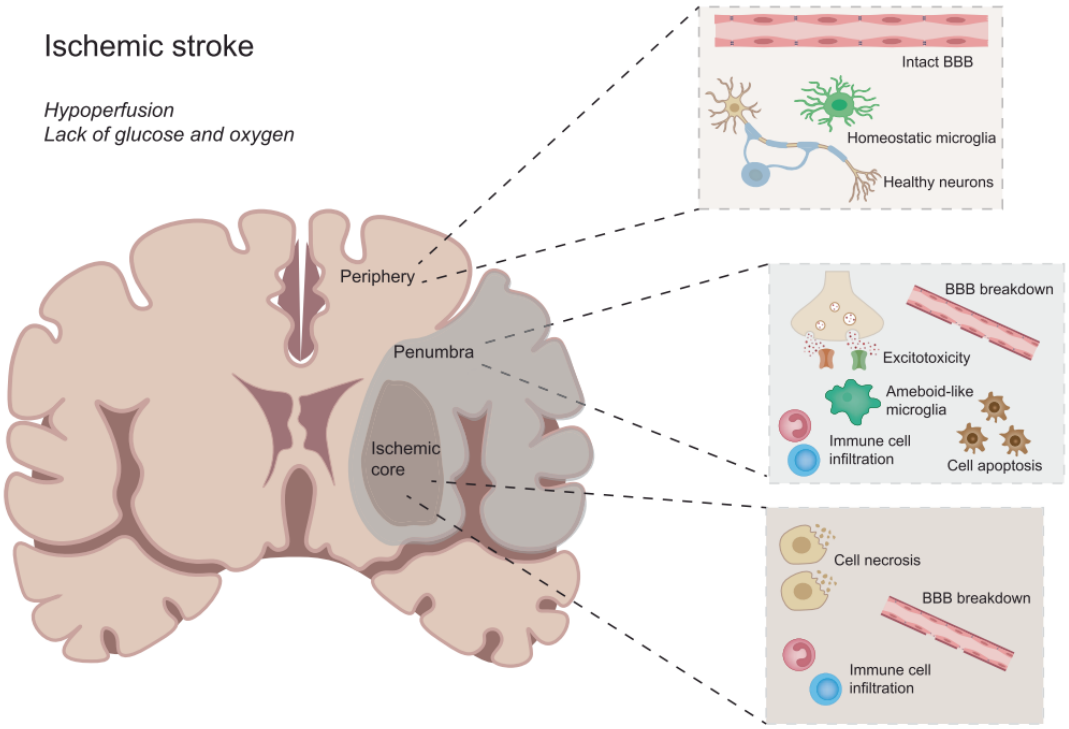

Figure 1.1: Illustration of an ischaemic stroke. Adapted from [2].

1.1.4 Treatment and recovery

Stroke management can be divided into two parts.

The first part aims to contain the stroke before it causes too much damage. As explained, the damage caused by a stroke is generally irreparable at the level of the ischaemic core. The challenge is therefore to rapidly restore perfusion in the region to supply the cells with oxygen and glucose. Depending on the cause of the ischaemia, blood flow is generally restored by treatment with drugs (thrombolytics, anti-coagulants, blood pressure regulators, and so on). In more extreme cases, it may be necessary to resort to surgery such as thrombectomy or Carotid endarterectomy [11]. Aside from the revascularization of the cells, it is also necessary to control the inflammation in the penumbra in order to halt the cell death processes, again through the use of appropriate drugs [1, 10].

On the other hand, rehabilitation treatment is also essential to the recovery of the patient's neuronal capacity. It also enables them to overcome possible long- and mid-term disabilities. These treatments are adaptive and tailored to the needs of individual patients. In all cases, rehabilitation should begin as soon as possible after the stroke to achieve the best results [12]. Rehabilitation treatments allow nerve connections to be reorganised through neuronal plasticity. These rearrangements result in the strengthening of healthy brain areas, to compensate for those that have been damaged, as well as changes to neuronal tracts and the establishment of new synaptic connections [13, 14]. To a lesser extent, and through phenomena that are still largely unknown, we can also add the creation of new neurons (neurogenesis) as a factor in recovery after stroke [7].

Rehabilitation treatments require monitoring methods that attempt to quantify the patient's psychomotor performance. In this way, it is possible to follow the patient's progress in order to adapt the treatment to his or her needs and to certify the effectiveness of the treatment. Various tests have been developed for this purpose. These include the Action Reach Arm Test (ARAT), which is used to assess the upper limb function [15, 16], the Stroke Impairment Assessment Set (SIAS), which is designed to highlight impairments compared to the unaffected side of the patient [17], and the 6-Minute Walk Test (6MWT). The latter is used in this study as a reference for the clinical measurement of the patients' condition. Its purpose is to measure the patient's endurance. In this test, the patient is asked to walk as far as possible in 6 minutes. The distance covered is used as a measure to evaluate the patient [18].

1.2 Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is a modern imaging technique, mainly used for virtual histology. Based on the magnetic properties of nuclei naturally present in our body, it uses the nuclear magnetic resonance (NMR) principles. The MRI has the advantages of being a relatively fast, cheap and automated technique compared to classical histology [19]. It is a non-invasive and non-destructive procedure, which makes it safer than techniques such as computed tomography (CT) or positron emission tomography (PET) relying on ionizing radiations.

1.2.1 Nuclear Magnetic Resonance principles

Atom's nuclei are composed of protons and neutrons (nucleons). Each of those particles rotates around an axis passing through its centre. As described by Figure 1.2, this rotation induces an angular momentum called "spin". With it comes an induced magnetic moment because of the structure of the nucleons, composed of charged particles: quarks. As a result, nuclei themselves have similar properties. Though a nucleus's spin and magnetic moment may not be trivial to estimate [20].

The idea lying behind the nuclear magnetic resonance (NMR) is to use nuclei's previously described magnetic properties and observe their magnetization changes under the influence of a strong, constant magnetic field and the excitations of radio-frequency pulses (RF pulses).

Because of their resulting magnetic moment, nuclei will behave like magnetic dipoles. If their orientations are randomly distributed at rest, when immersed in a constant magnetic field $\boldsymbol{B_0}$, their spinning direction will tend to align. This alignment depends on the spin of the considered nucleus and takes discrete positions due to quantum effects. Those positions are not perfectly aligned with the magnetic field streamlines and induce a precession of the nuclei around the magnetic field direction (see Figure 1.3). The rotation occurs at a frequency called the Larmor frequency ν_0 which is given by the Larmor equation :

$$\nu_0 = \frac{\gamma B_0}{2\pi}, \tag{1}$$

where γ is the gyromagnetic ratio of the nuclei. This ratio and the alignment's possibilities of the nuclei depend on their nature and spin number respectively. From now on, the ^{1}H isotope whose spin S equals $\frac{1}{2}$ will be considered for future development as it is the target of the diffusion magnetic resonance imaging (dMRI) technique.

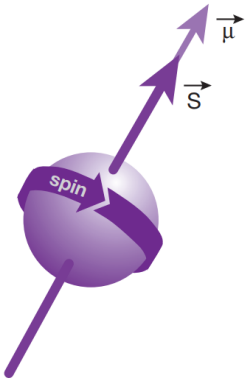

Figure 1.2: Illustration of a rotating particle with its spin ($\vec{S}$) and magnetic moment ($\vec{\mu}$) [21]

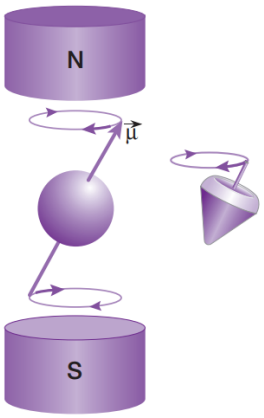

Figure 1.3: Illustration of a rotating particle precessing in a magnetic field [21]

For $S = \frac{1}{2}$ nuclei such as ^{1}H, the repartition of the spin orientations is split between two spin states: $m_S = \frac{1}{2}$, called parallel (and noted $\uparrow$) as it is oriented along the direction of the $\boldsymbol{B_0}$ magnetic field, and $m_S = \frac{-1}{2}$, called anti-parallel (noted $\downarrow$). The parallel and anti-parallel spins correspond to two different energy levels of the particles:

$$E_\uparrow = \frac{\gamma\hbar B_0}{2} \qquad \text{and} \qquad E_\downarrow = -\frac{\gamma\hbar B_0}{2}, \qquad \text{with} \qquad \Delta E = \gamma\hbar B_0 = \nu_0 h. \tag{2}$$

At equilibrium, the repartition of the particles among those energy levels is not perfectly symmetric and the asymmetry of the repartition $\frac{N_\uparrow}{N_\downarrow}$ is given by the Boltzmann distribution:

$$\frac{N_\uparrow}{N_\downarrow} = \exp\left(\frac{\Delta E}{kT}\right), \tag{3}$$

N denoting the number of particles in the considered state, k the Boltzmann's constant and T referring to the temperature in Kelvin. This asymmetry, although slight[1], leads to a non-zero resulting magnetic moment when summing the contribution of each particle in a region. This macroscopic magnetization $\boldsymbol{M}$ has a norm $M_0 = \frac{\gamma^2\hbar^2\rho B_0}{4kT}$, where ρ stands for the spin density (i.e. the number of considered particles per unit volume), and points toward the exact same direction than $\boldsymbol{B_0}$.

This magnetization being very weak compared to the surrounding magnetic field, it is impossible to measure directly. Yet, it is possible to create conditions in which a measure is possible by perturbing the system (Figure 1.4). As explained above, alignments of the spins correspond to the energy levels of their respective nuclei. In quantum mechanics, it is possible to raise the energy of a particle from one state to the other by giving it the right amount of energy in the form of a quantum. The energy needed is given by the difference between the two energy levels (see Equation 2), meaning that applying radio frequency pulses at Larmor frequency to the nuclei will make them "flip" from the parallel state to the anti-parallel one. The RF-pulse is applied as an oscillating magnetic field $\boldsymbol{B_1}$ in a perpendicular direction of $\boldsymbol{B_0}$. On top of

[1] $\frac{N_\uparrow}{N_\downarrow} = 2ppm$ at body temperature and with $B_0 = 0.5T$ according to [21]

the change in the energy state imputed to the particles, this excitation also rephases them together. The global effect of such an interaction can be seen as a rotation of the macroscopic magnetization $\boldsymbol{M}$ around the direction of $\boldsymbol{B_1}$. The angle α $\boldsymbol{M}$ travels during the pulse depends both on the magnitude of $\boldsymbol{B_1}$ and the duration of the pulse (see Figure 1.5). Its name is "the flip angle" and can be fully controlled by acting on the previously mentioned parameters. As $\boldsymbol{M}$ rotates around $\boldsymbol{B_1}$, it can be decomposed in 2 components $\boldsymbol{M_z}$, the longitudinal magnetization, parallel to the main magnetic field $\boldsymbol{B_0}$, and the rotating $\boldsymbol{M_{xy}}$, perpendicular to $\boldsymbol{B_0}$. As its direction is perpendicular to the one of $\boldsymbol{B_0}$, $\boldsymbol{M_{xy}}$ allows us to significantly measure a signal apart from the overwhelming magnitude of $\boldsymbol{B_0}$ [21, 22, 23].

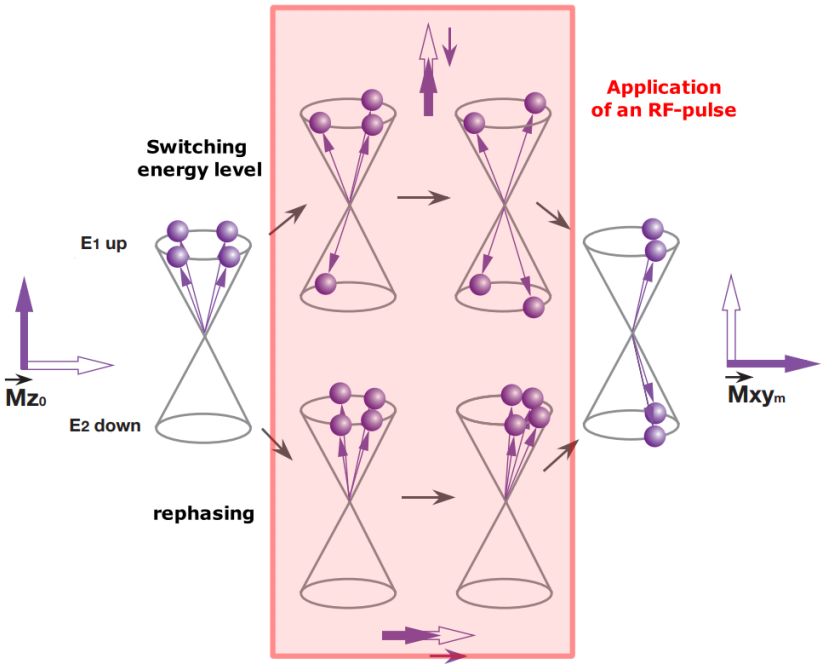

Figure 1.4: Illustration of the application of an RF-pulse. Adaptated from [21].

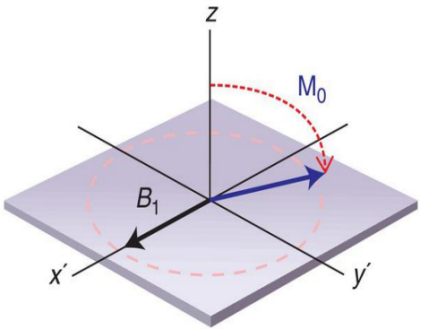

Figure 1.5: Representation of the flip angle. [23]

1.2.2 Relaxation types

The excited state in which the particles are left after the application of an RF-pulse is unstable and they will recover their equilibrium through two processes of relaxation.

T1-relaxation

The first relaxation process, referred to as T1-relaxation, is called longitudinal relaxation as it concerns the recovery of the net magnetization M_z to its original value M_0. It is also called the spin-lattice relaxation in reference to the interaction inducing the phenomenon. It corresponds to the de-excitation of the spins progressively getting back to their original energy level by restituting their energy excess to the surrounding lattice as a heat exchange. The relaxation begins as soon as the RF-pulse stops and the time evolution of M_z is described by an exponential returning to M_0 with a characteristic time T_1 as described by formula (4):

$$M_z(t) = \exp\left(\frac{-t}{T_1}\right)[M_z(0) - M_0] + M_0. \tag{4}$$

After a time T_1, the longitudinal magnetization recovers 63% of its nominal value, after $2T_1$ it is 87%, $3T_1$ 95% and so on. For biological tissues, T_1 usually gets values between 500 and 1000 milliseconds, depending on the molecular properties of the surrounding matter and its state. More precisely, it is the molecular collision frequency of the medium that determines the relaxation time. The more this frequency is close to the Larmor frequency of the spins, the easier it is for spins to transfer their energy surplus to the environment and smaller is T_1 [21, 22, 23].

T2-relaxation

The second relaxation process, referred to as T2-relaxation, is called transverse relaxation and concerns the loss of signal in M_{xy} due to the dephasing of the spins. During the RF-pulse, all spins tend to precess at the exact same frequency and gain coherence in their precession as they are rephasing. As soon as the RF-pulse stops, the slight differences in the biological medium occupied by the spins induce small inhomogeneities in the magnetic field $\boldsymbol{B_0}$. These differences in the field shift the local Larmor frequencies and spins precess at different speeds, quickly losing their coherence. These inhomogeneities in the global magnetic field result from spin-spin interactions inducing microscopic local magnetic fields superposing with B_0. This interaction gives its name to the T2-relaxation as it is also called "spin-spin relaxation". In contrary with T1-relaxation, T2-relaxation does not involve an energy exchange between the spins and their surroundings.

The decrease in the transverse magnetisation M_{xy} through time is given by formula (5) which describes an exponential returning to zero with a characteristic time T_2.

$$M_{xy}(t) = M_{xy}(0)\exp\left(\frac{-t}{T_2}\right). \tag{5}$$

In this case, waiting a time T_2, $2T_2$ and $3T_2$ lead to a loss of the M_{xy} magnetisation of 63%, 87% and 95% respectively. For biological tissues, T_2 values usually lie between 50 and 100 milliseconds depending on the molecular properties of the medium and its state of matter. Note

that those values for T_2 are much smaller than the ones obtained for T_1.

In practice, though, it appears that the decay time T_2 is much smaller than expected similar to what is shown in Figure 1.6. As explained, the loss of coherence between precessing spins is a consequence of inhomogeneities in the surrounding magnetic field. These inhomogeneities are in fact the combination of perturbations due to molecular interactions (that is where lies our interest) and irregularities of the artificial magnetic field produced by the imaging device. Such irregularities have a massive impact on the dephasing of the spins and accelerate the T2-relaxation. The exponential decay is thus effectively characterized by a smaller time $T_2^* < T_2$. The inhomogeneities imputed to the device's irregularities have the advantage of remaining constant during the acquisition. It is thus possible to cancel their effect and there exist several ways to proceed. Among them, the use of specific imaging sequences using the so-called "echo" effect or the acquisition of a "blank image" without the patient to map the field inhomogeneities [21, 22, 23].

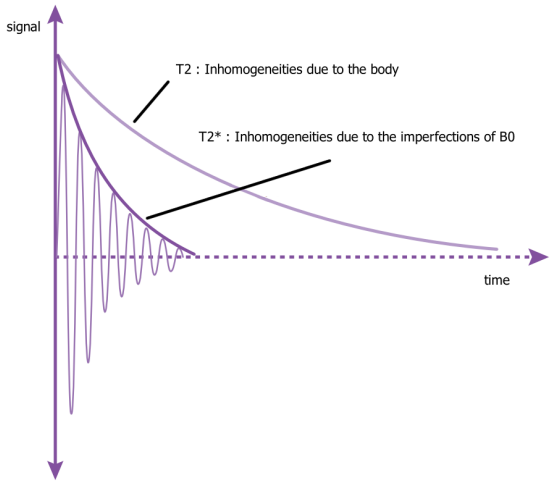

Figure 1.6: Illustration of the T_2^* effet. Adapted from [21]

1.2.3 Free Induction Decay

As explained above, the valuable signal in an MRI comes from the transverse magnetization $\boldsymbol{M_{xy}}$ whose orthogonality with $\boldsymbol{B_0}$ makes it distinguishable. The signal induced by the oscillating evolution of $\boldsymbol{M_{xy}}$ is called the Free Induction Decay (FID) signal. The signal is received by a coil in the device using Faraday's law of induction which states that any change in a magnetic flux through a closed loop induces a proportional electromotive force in that same loop. As a result, the FID signal is converted into a measurable electronic signal. This signal takes the shape of a sinusoid contained in an exponentially decreasing envelope as depicted in Figure 1.7. The oscillation of the curve is due to the rotation of $\boldsymbol{M_{xy}}$ around $\boldsymbol{B_0}$ and the exponential decay describes the T2-relaxation.

The FID signal depends on three biological factors of the tissues :

1. The T_1 relaxation time

2. The T_2 relaxation time

3. ρ, the protonic density

Without entering into the details, it is good to know that modifying the parameters of the used acquisition sequence modifies the influence of those features on the signal. The obtained image thereby provides completely different information about the visualized tissues depending on the contrast that has been chosen (see Figure 1.8) [21, 22, 23].

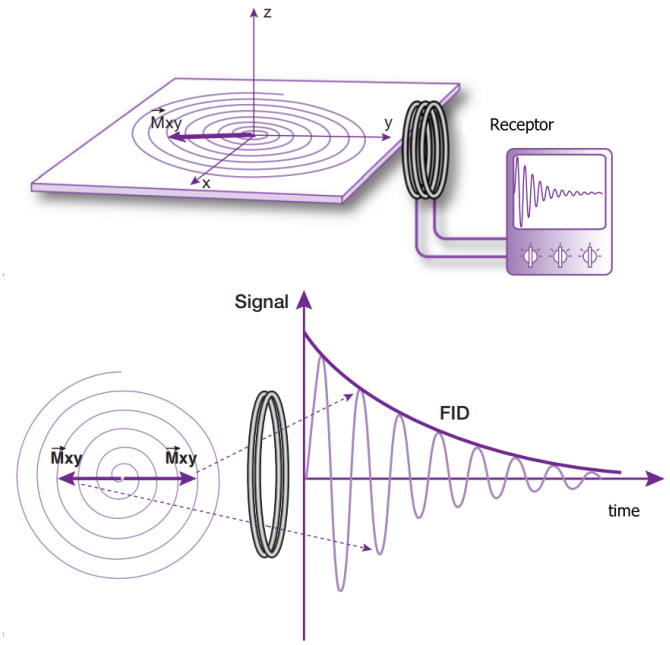

Figure 1.7: What is the FID and how it is intercepted. Adapted from [21]

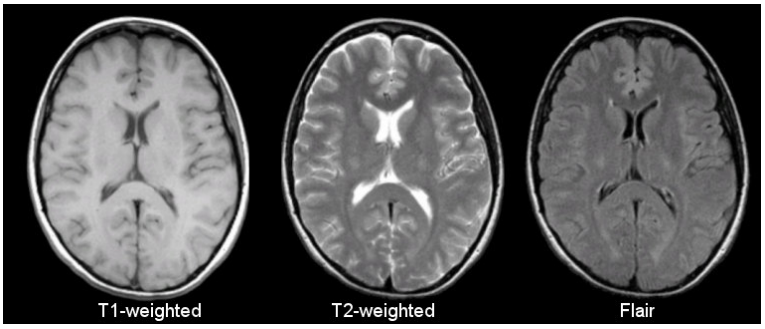

Figure 1.8: Differences between T1, T2, and flair (protonic density) contrasts. [24]

1.2.4 Spatial encoding

From the FID signal, the MRI has to rebuild an entire three-dimensional image of the subject. It is thus necessary to encode the needed spatial information into the signal. A three-dimensional image is constituted of voxels, the generalization of pixels to 3 dimensions. Each voxel therefore has to be localized along three directions. To locate a voxel along the $\boldsymbol{z}$ direction (which

corresponds to the direction of $\boldsymbol{B_0}$), a gradient $\boldsymbol{G_{ss}}$ can be applied to the main magnetic field. Indeed, since the excitation of spins relies on the Larmor frequency and since, according to equation 1, this frequency depends on the surrounding magnetic field intensity, applying a gradient while providing the RF-pulses allows to perform a slice selection in the subject. By changing the frequency window of the RF-pulses, it is then possible to choose which slice is affected by the pulses. The width of those slices depends on the width of the chosen window (see Figure 1.9). The total image is constructed slice by slice.

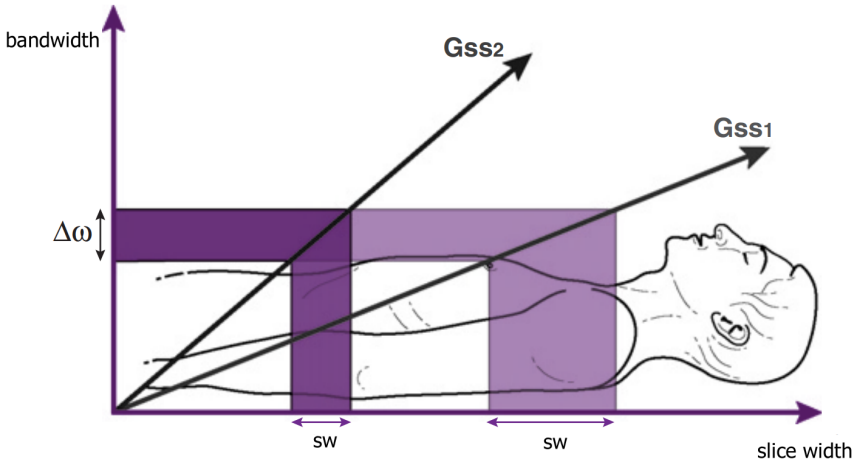

Figure 1.9: Relationship between the RF-pulse bandwidth and the width of the selected slice depending on the chosen gradient. Adapted from [21].

Once the slice is chosen, the voxel still has to be found in the $\boldsymbol{x}$-$\boldsymbol{y}$ plane (perpendicular to $\boldsymbol{B_0}$). The same method cannot be applied to locate the signal on that plane. That is why, MRI uses phase and frequency encodings. In fact, instead of directly filling the image slices with voxels, the MRI fills the Fourier space (also named k-space) of the slices. The image is then rebuilt using a 2D Fourier transform. The way the k-space is filled depends on the MRI sequence used but the principle always remains the same. In the k-space, one axis is coded by the mean of frequencies and the other by the mean of phases, both relying on magnetic gradients applied along different directions and at different times in the sequence.

The frequency encoding is constructed using a magnetic gradient $\boldsymbol{G_\omega}$ applied in a transverse direction (in the $\boldsymbol{x}$-$\boldsymbol{y}$ plane) during the read-out of the FID signal. This gradient affects the precession speed of $\boldsymbol{M_{xy}}$ around the $\boldsymbol{z}$-axis, altering the received signal. Therefore, instead of a single, exponentially damped sinusoidal curve, the signal is a combination of different frequencies sinusoids in an exponentially decreasing envelope as depicted in Figure 1.10.

The phase encoding is obtained using a magnetic gradient $\boldsymbol{G_\phi}$ in the last remaining direction of the space (always in the $\boldsymbol{x}$-$\boldsymbol{y}$ plane). This last gradient is applied before the reading of the signal, outside the RF pulse application windows. In the presence of $\boldsymbol{G_\phi}$, the spins rotate at different speeds in the gradient direction, inducing a phase shift of the spins in the same direction. Because the spins precess at the same speed in the absence of any gradient, this shift is conserved until the signal is read as shown in Figure 1.11 [21, 22, 23].

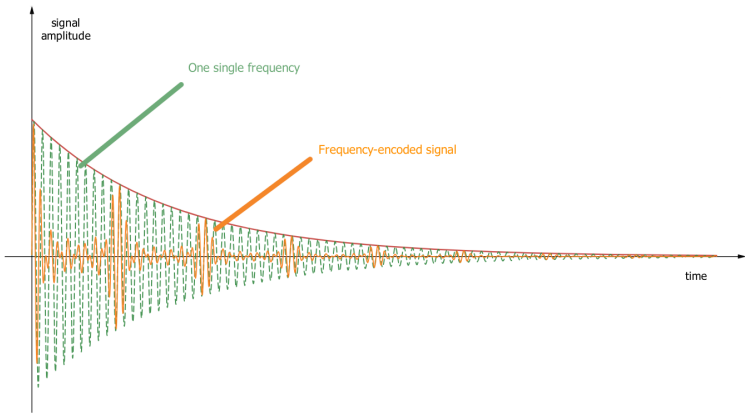

Figure 1.10: Shape of a frequency-encoded signal.

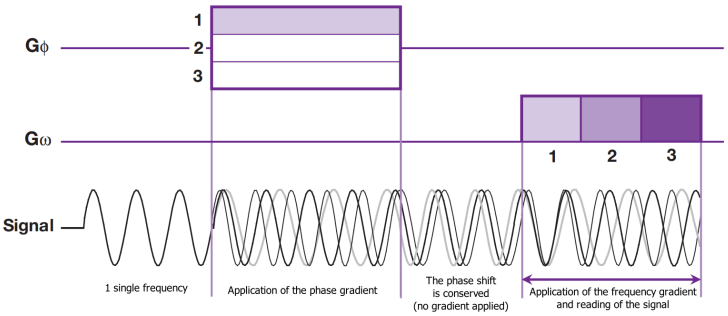

Figure 1.11: Step-by-step spatial encoding of the signal in one slice. Adapted from [21].

1.2.5 Magnetic Resonance Imaging artefacts

Magnetic resonance imaging is of course subject to artefacts, phenomenons that can alter the quality of the provided images. Examining those artefacts allows one to understand their impact on the images in order to, either correct them if possible or be critical about a diagnosis. It is possible to draw up a non-exhaustive list of these artefacts.

Motion artefacts

MRI images intrinsically require time to be acquired. Even though some sequences may accelerate the acquisition process, the time needed is usually sufficiently long to allow the patient to move during the imaging. The patient's movements may be random such as physical movements and produce smearing in the phase encoding direction of the image. Periodic movements also occur in a wide enough time window. These are cardiac or pulmonary movements. Periodic movements usually produce a ghosting effect in the images [25, 26, 27] (see Figure 1.12). Reducing the effects of this kind of artefact can be achieved by using fast imaging techniques, movement correction techniques or by trying to follow the motion during the acquisition [26].

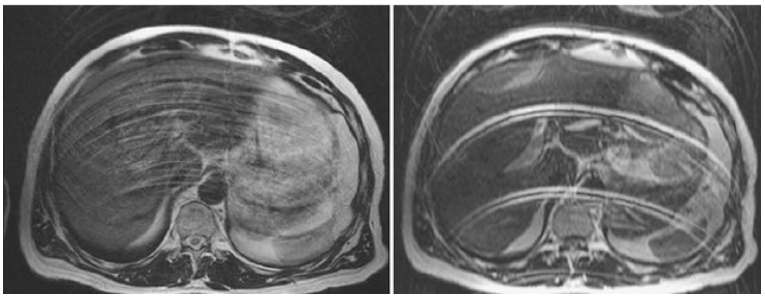

Figure 1.12: Motion artifact. More random respiratory motion (left) shows more general blurring of the T2 weighted TSE image (TR 2200 ms, TE 103 ms, ETL 21), periodic breathing leads to ghost artifacts (right). [27]

k-space artefacts

Some artefacts are directly related to the fact that the MRI acquires the image in the k-space and then uses a Fourier transform to convert it into the usual space. One common artefact is the "aliasing" or "wrap-around" artefact. It appears when the field of view (FOV) is not properly adjusted to the subject and is smaller than required. In this case, the part of the subject that is outside the field of view is superimposed on the opposite side of the image as shown in Figure 1.13. To avoid this artefact, the FOV can be increased or the part of the subject outside of the subject can be saturated. Another k-space artefact is known as the "Gibbs ringing" and consists of spurious ringing around sharp edges in the image (see Figure 1.14). It is directly induced by the discreet nature of the MRI. Indeed, sharp edges cannot be perfectly reconstructed with a finite number of frequencies using the Fourier transform. We say that it is truncated. To avoid this effect, we may use a low-pass filter on the data which would induce a blurring effect along with a reduced spatial resolution, or we could increase the number of voxels while keeping a constant FOV [23].

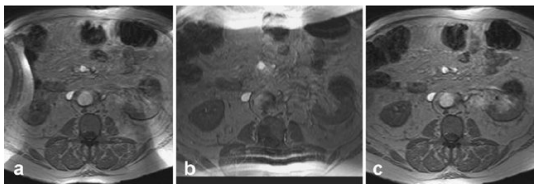

Figure 1.13: Wraparound artifact. a : Axial T1 weighted image (GRE, TR 80 ms, TE 6.9 ms) shows wraparound of the parts of the lateral body wall, which are outside of the FOV, in the left-right phase encoding direction. b : Repeat scan with identical parameters but after switch. [27]

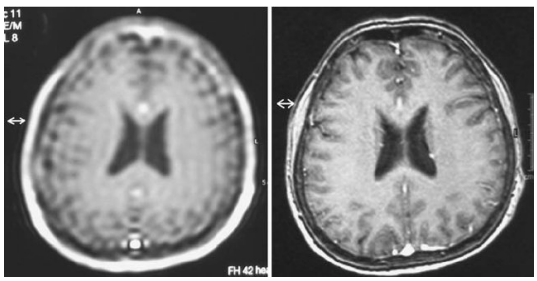

Figure 1.14: Truncation. Ringing artifact with bright and dark lines parallel to edges of abrupt intensity changes, in a T1 weighted SE image (TR 252 ms, TE 15 ms), with low matrix size (64×64). After increasing the matrix (256×256) the ringing artefact is not visible . [27]

Magnetic field perturbation artefacts

Other artefacts may be induced by some perturbations of the magnetic field during the MRI. These perturbations may come from different sources. Possible sources of perturbations are induced Eddy current in the tissues of the imaged body. Those arise when switching on and off the magnetic field gradients and may provoke distortions, shifts or signal losses in the image. Generally, most of the effects of Eddy currents are compensated directly by the manufacturer of the MRI device. In some cases though, the presence of an unexpected object may cause similar effects. For example, a metallic prosthesis, a tooth implant or anything else may have a significant impact on the resulting image as depicted in Figure 1.15. [23, 25, 26, 27]

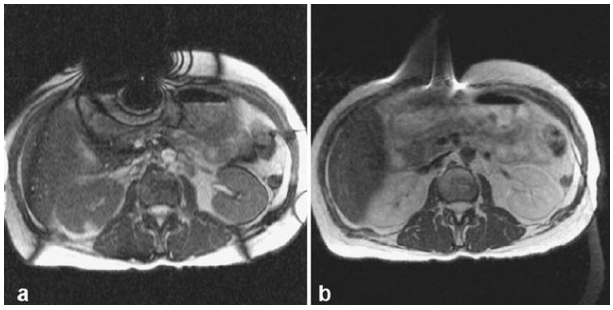

Figure 1.15: Susceptibility artifact. Two manifestations of a susceptibility artifact originating from a metallic part of a cathe- ter: a (SE, TR 11 ms, TE 5 ms) shows severe blooming around the ferromagnetic material, whereas b (SE, TR 4500 ms, TE 96 ms) shows bizarre distortion of the abdominal wall of the same patient. [27]

RF and gradient artefacts

Another factor that can cause artefact images is imperfections in the RF signals and in the magnetic gradients. We can mention nonlinearities in the gradients distorting the image or unwanted RF sources interfering with the receptor and inducing noise as well as Moiré fringes in the image (see Figure 1.16 and 1.17). [23, 25, 26, 27]

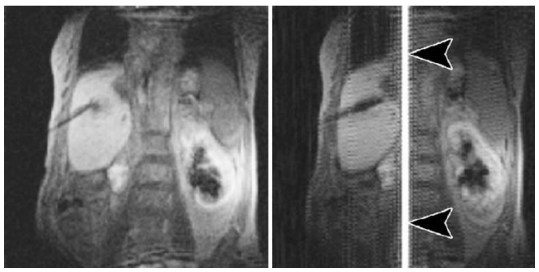

Figure 1.16: Zipper artifact. Coronal image (TR 53.7 ms, TE 7.4 ms) shows an RF ablation needle placed in the liver (left). After switching the radiofrequency on, zipper artifacts are produced (arrows, right image). [27]

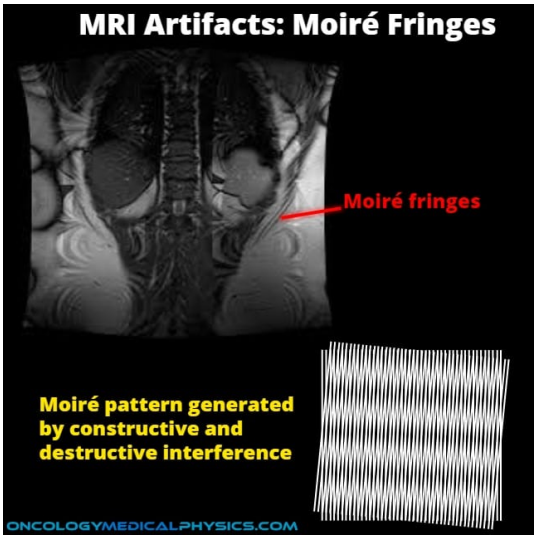

Figure 1.17: Example of Moiré fringes as an RMI artefact. [28]

Chemical shift artefacts

At last, inhomogeneities in body tissues also induce artefacts. These are known as "chemical shifts". They are the result of the differences in the chemical properties of the different tissues making up the human body [23, 25, 26, 27]. For example, as the chemical composition of water and fat differ from one another, their resonance frequency is also different. This leads to misregistration of the considered molecules. A shift can be observed between the information acquired from water molecules and fat molecules. Usually most prominent along the phase-encoded direction, it may draw brighter and darker bands at the border of different tissues in the resulting image. (see Figure 1.18).

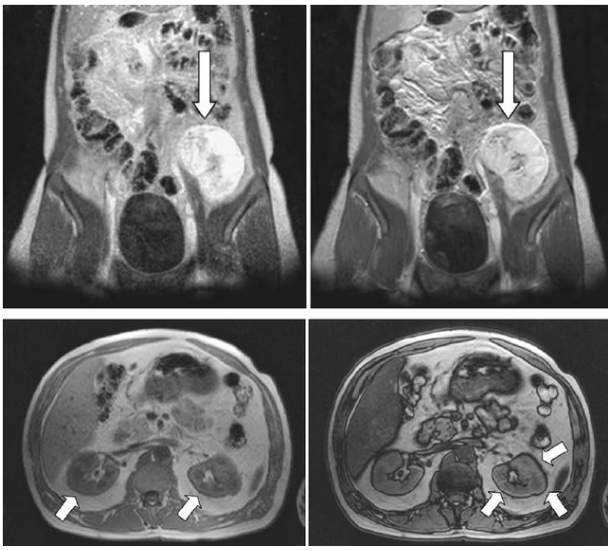

Figure 1.18: Chemical shift artifact of the first and second kind. Chemical shift of the first kind (left) on the T1w in-phase GRE image (TR 100 ms, TE 4.4 ms) shows a dark rim in the posterior kidney-fat interface. Frequency encoding direction is anterior- posterior. Chemical-shift artifact of the second kind: T1w opposed-phase GRE image (right) (TR 100 ms, TE 2.2 ms) shows a surrounding dark rim at the fat- tissue interface, independent of the frequency encoding direction. [27]

1.3 diffusion Magnetic Resonance Imaging

Diffusion Magnetic Resonance Imaging (dMRI) or Diffusion-Weighted Magnetic Resonance Imaging (DW-MRI) is another form of imagery based on the same principle of magnetic resonance. Its purpose is to capture the water molecule diffusive movements in the body. Because the diffusion of water molecules is influenced by the structure of the environment, obtaining information about the diffusion of these molecules gives us indirect information about the layout of the tissues observed. Indeed, the unconstrained diffusion of water molecules follows a Brownian motion. In the body, and the brain in particular, this diffusion can be constrained by natural barriers such as cell and axonal walls. This reduces water diffusion in the constrained directions. Depending on the layout of the tissues, diffusion may be isotropic, anisotropic, relatively strong or relatively weak. Therefore, water molecules evolving in CSF will tend to diffuse strongly and isotropically, whereas water molecules present in axons follow a strongly anisotropic diffusion aligned with the direction of the axon. It is thus common practice in diffusion imaging to inspect the diffusivity of water in several directions in order to extract the maximum amount of information about the cerebral structure. [4, 5, 6, 21]

To obtain a diffusion-weighted image, we label the molecules using their phase. In the same way as for phase encoding, the application of a gradient in the magnetic field dephases the spins. Assuming that the water molecules are perfectly static, applying the same but opposite gradient after a certain lapse of time would put all the spins in phase again. However, water molecules diffuse. Therefore, if a water molecule moves in line with the gradient, applying the opposite gradient will not perfectly compensate for the phase shift. The associated spin therefore remains out of phase. It should be remembered that the signal read by the receiver comes from the $\boldsymbol{M_{xy}}$ vector. The norm of $\boldsymbol{M_{xy}}$ depends exclusively on the phase match of the spins in a voxel. If some water molecules have scattered and become out of phase, their contribution to the signal is reduced. In this way, a voxel whose diffusion is relatively weak in the direction of the gradient produces a strong signal (brighter in the image) and, conversely, a voxel whose diffusion is relatively strong in the direction of the gradient produces a weak signal (darker in the image) as it can be seen in Figure 1.19. the operation can be repeated by changing the direction of the gradient in order to obtain diffusion of the water molecules in that same direction [4, 5, 6, 21].

In addition to the direction analysed, diffusion imaging is affected by several parameters: the intensity of the gradient (G), the duration of gradient application (δ), the gyromagnetic ratio (γ) and the time between application of the gradient and its opposite (Δ). All these parameters contribute to the *b-value* which characterizes the acquisition sequence and is defined as :

$$b = (\gamma G \delta)^2 (\Delta - \frac{\delta}{3}). \tag{6}$$

This value (in s/mm^2) can be used to estimate the signal intensity in the considered direction. It is described by the Stejskal-Tanner equation :

$$S(b) = S(0) exp(-bD), \tag{7}$$

with $S(b)$ the signal value for the *b-value* b and D the diffusion constant specific to the tissue [21, 29].

From here, we can introduce a new quantity: the apparent diffusion coefficient (ADC), which, in its own way, gives an account of diffusion phenomena. The ADC (in mm^2/s) is given by the following formula:

$$ADC = -log(\frac{S}{S_0})/b. \tag{8}$$

For a given direction, it indicates how quickly the diffusion signal ($S(b)$) will decrease as a function of b. A small ADC indicates strongly constrained diffusion and a large ADC indicates free diffusion (see Figure 1.19) [21].

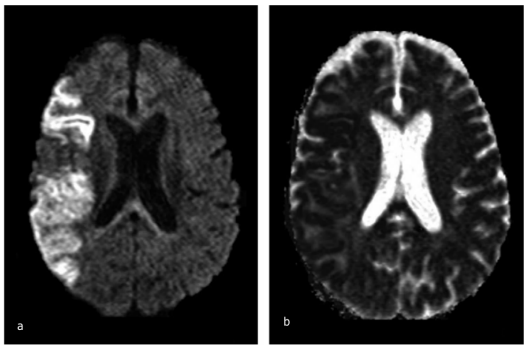

Figure 1.19: Diffusion-Weighed image (b-*value* = 1000) after an ischaemic stroke (a). ADC map of the same image (b). Adapted from [21]

1.4 Microstructural models

As explained above, diffusion phenomena are intrinsically linked to the configuration of biological tissues. Diffusion imaging therefore enables us to build models that attempt to reconstruct the tissue composition of the brain and its microstructures. It is clear that the resolution of an MRI image (1 or $2mm$) does not allow microstructures to be accurately identified (a few μm). In this section, we nevertheless show a few models that can provide us with information about the brain's microstructures.

1.4.1 DTI

Diffusion Tensor Imaging or DTI is the most basic microstructural model. It consists of generating and analysing a second-order tensor representative of the diffusion within a voxel (see Equation 9).

$$\boldsymbol{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix}. \tag{9}$$

The diffusion tensor is made up of 9 elements, each corresponding to the amplitude of the diffusion coefficient in a given direction. To fill this tensor, it is therefore necessary to obtain a DW-MRI along 9 non-collinear directions.

In fact, a 3x3 tensor such as this represents an ellipsoid characteristic of diffusion in the voxel (as shown in Figure 1.20). By analysing the 3 eigenvectors and the 3 eigenvalues (λ_1, λ_2, λ_3) of this tensor, we can in fact determine the preferred directions of diffusion and the associated diffusion coefficient. This is what DTI is all about: finding the tensor's eigenvectors and their associated eigenvalues. The DTI also suppose that the tensor is symmetric reducing the number of distinct components to 6 [30, 31, 32, 33].

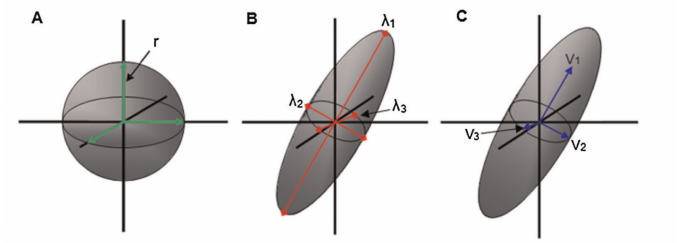

Figure 1.20: Isotropic diffusion (A); Anisotropic diffusion and eigenvalues (B); Anisotropic diffusion and eigenvectors (C). Adapted from [32].

From the diffusion tensor, different metrics have been established to characterize the microstructure in the voxels.

Mean Diffusivity (MD)

MD is, as its name suggests, the diffusivity averaged over all the main directions of diffusivity. In other words, it is the average of the eigenvalues of the $\boldsymbol{D}$ tensor.

$$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}. \tag{10}$$

A high MD is therefore generally characteristic of unrestricted diffusion zones: CSF, oedema, necrotic areas, etc. Note that it does not take into account the direction of distribution [29, 34].

Axial Diffusivity (AD)

The AD is simply the first eigenvalue of the diffusion tensor:

$$AD = \lambda_1. \tag{11}$$

It is the value of the diffusion coefficient in the dominant diffusion direction of the voxel. A large AD can represent a strong tendency for axons to follow a certain direction. On the other hand, a decrease in the value of the AD can indicate a deterioration of the axons [29, 34].

Radial Diffusivity (RD)

RD, similar to AD, represents the diffusivity in the plane perpendicular to the voxel's dominant diffusion direction. It is the average of the remaining eigenvalues.

$$RD = \frac{\lambda_1 + \lambda_2}{2}. \tag{12}$$

The RD is often dependent on the myelination of axons or the thickness of an axon bundle [29, 34].

Fractional Anisotropy (FA)

FA measures the fraction of the "magnitude" of $\boldsymbol{D}$ attributed to its anisotropic character. Thus, a voxel with an FA close to 1 would correspond to an area of the brain in which water diffuses in a highly anisotropic way. A region containing many aligned axons, for example. Conversely, an FA of almost zero suggests that diffusion is isotropic. More roughly, it is an indicator of the shape of the ellipse associated with the tensor. The smaller the FA, the closer the ellipse is to a sphere. The larger the FA, the more elongated the ellipse. Its formula is given by:

$$FA = \sqrt{\frac{3}{2}}\sqrt{\frac{(\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}. \tag{13}$$

[29, 34, 35]

1.4.2 NODDI

The main advantage of the DTI model is its simplicity. However, it has a price. The DTI metrics are extremely limited by the choice of representing the whole volume of a voxel by one single 3x3 tensor. The Neurite Orientation Dispersion and Density Imaging (NODDI) model introduces the idea of a multi-compartment model, no longer considering the voxel as a homogeneous volume but as the union of subspaces with different properties. It proposes to describe A the total diffusion in the voxel as the weighted sum of the contributions from each compartment, leading to the following expression :

$$A = (1 - \nu_{iso}) * (\nu_{ic} A_{ic} + (1 - \nu_{ic}) A_{ec}) + \nu_{iso} A_{iso}, \tag{14}$$

where A_{ic}, A_{ec}, A_{iso} stand for normalized signals of an intracellular, an extracellular and a CSF compartment respectively, and ν_{ic}, ν_{ec}, ν_{iso} designate the volume fractions of the same compartments [36].

Intracellular compartment

It represents the space occupied by neurites. Because the cellular boundaries are shaped like tubes in neurites, the diffusion in this space is supposed to be highly restricted and anisotropic. In the model, it is considered that the diffusion is completely hindered perpendicularly to the axons' axis. This is represented, by a set of sticks (cylinders with a null radius) centred on the voxel and more or less aligned. The normalized signal emitted by the compartment is given by :

$$A_{ic} = \int_{\mathbb{S}^2} f(\boldsymbol{n}) \exp(-b d_{||} (\boldsymbol{q} \cdot \boldsymbol{n})^2) d\boldsymbol{n}, \tag{15}$$

where the exponential describes the signal attenuation in a stick of intrinsic diffusivity $d_{||}$ and oriented along $\boldsymbol{n}$, $\boldsymbol{q}$ and b stand for the gradient direction and the *b-value* of the sequence respectively, and $f(\boldsymbol{n})d\boldsymbol{n}$ gives the probability to find sticks along the orientation $\boldsymbol{n}$.

Furthermore, f is modelled according to a Watson distribution involving :

$$f(\boldsymbol{n}) = M(\frac{1}{2}, \frac{3}{2}, \kappa)^{-1} \exp(\kappa(\boldsymbol{\mu} \cdot \boldsymbol{n})^2), \tag{16}$$

with M is a confluent hypergeometric function, $\boldsymbol{\mu}$ stands for the mean orientation of the neurites and κ is a concentration parameter that measures the extent of orientation dispersion about the mean [36].

Extracellular compartment

This compartment designates the space in the periphery of neurites. In this space lies various glial cells and other bodies constraining the diffusion of water molecules and leading to anisotropy in the diffusion. Yet we consider here that the restriction on the Brownian motion is not sufficient to be fully constrained. Therefore, the diffusion is modelled by an anisotropic Gaussian distribution in the following equation for the contribution of the compartment to the normalized total signal :

$$A_{ec} = \exp\left(-b\boldsymbol{q}^T \left(\int_{\mathbb{S}^2} f(\boldsymbol{n}) D(\boldsymbol{n}) d\boldsymbol{n}\right) \boldsymbol{q}\right), \tag{17}$$

with $D(\boldsymbol{n})$ a cylindrically symmetric tensor having a principal direction of diffusion $\boldsymbol{n}$, diffusion coefficients $d_{||}$ and $d_{\perp}$ parallel and perpendicular to $\boldsymbol{n}$ respectively. Whereas the parallel diffusivity is considered to be the very same as the intrinsic free diffusivity of the intracellular subdivision, the perpendicular diffusivity is said to obey a simple tortuosity model [36]:

$$d_{\perp} = d_{||}(1 - \nu_{ic}). \tag{18}$$

Cerebrospinal fluid compartment

The CSF compartment represents the remaining space in the voxel, free of neurites and other cells. The diffusion here is modelled as a pure unconstrained isotropic Gaussian distribution. This leads to a contribution to the normalized signal that follows the relation :

$$A_{iso} = \exp(-b d_{iso}), \tag{19}$$

with d_{iso} the isotropic diffusion coefficient of the medium [36].

Benefits and limitations

The NODDI model allows us to reach interesting properties of the microstructure in a voxel. The heterogeneity which is brought by the model gives the possibility to investigate beyond the basic metrics provided by DTI. It is interesting to analyse the volume repartition between the different compartments inside the voxels, telling us something about the state of health of the brain in the considered area. We may also derive the fibre Orientation Dispersion Index (ODI) of the voxels :

$$ODI = \frac{2}{\pi}\arctan\frac{1}{\kappa}, \tag{20}$$

which can be useful to estimate the variability in the neurite orientations [36].

On the other hand, NODDI is far from a perfect model and has its drawbacks. In particular, its simplifying assumptions, which, even though facilitating the fitting of the model, can lead to errors in the estimation of certain parameters. Similarly, the model itself can sometimes be too superficial to correctly capture all the complexity of brain tissue.

1.4.3 DIAMOND

The DIstribution of Anisotropic MicrOstructural eNvironments in Diffusion compartment imaging (DIAMOND) model keeps the idea of a multi-compartments model, trying to improve its representativeness with regard to the biological reality of brain tissues. Whereas NODDI relies on underlying assumptions about tissues and their effects on the diffusion detected, DIAMOND opts for a statistical distinction between compartments. It considers that each voxel is an amalgam of heterogeneous populations of tissues, themselves subject to heterogeneity. In other words, within a single voxel there are several categories of tissue with different diffusive properties, each category of tissue can be likened to a distinct three-dimensional diffusion distribution of its associated spins. Each spin packet then constitutes a compartment in the model. Within a compartment, the distribution of spins is modelled by a three-dimensional anisotropic Gaussian distribution as seen in Figure 1.21. Each of these distributions is therefore characterised by its own diffusion tensor $\boldsymbol{D}$. The total signal received during imaging (S_k) is then given by the following formula:

$$S_k = S_0 \int_{\boldsymbol{D}\in Sym^+(3)} P(\boldsymbol{D})\exp(-b_k \boldsymbol{g}_k^T \boldsymbol{D} \boldsymbol{g}_k^T) d\boldsymbol{D}, \tag{21}$$

with S_0 is the signal obtained for a b_0 *b-value*, $Sym^+(3)$ represents the set of 3×3 symmetric positive-definite matrices, $\boldsymbol{g}_k$ is the diffusion gradient and b_k its associated *b-value*. $\mathrm{P}(\boldsymbol{D})$ designates here a matrix-variate distribution that represents the fraction of spin packets having the diffusion tensor $\boldsymbol{D}$. In order to best represent the reality of the distribution of spins within a packet, the model chooses to use the peak-shapped mv-Γ distribution. This distribution is given, for any matrix $\boldsymbol{D} \in Sym^+(p)$, by:

$$P_{\kappa,\Sigma}(\boldsymbol{D}) = \frac{|\boldsymbol{D}|^{\kappa-(p+1)/2}}{|\Sigma|^{\kappa}\Gamma_p(\kappa)} \exp\left(-\mathrm{Tr}(\boldsymbol{\Sigma}^{-1}\boldsymbol{D})\right), \tag{22}$$

with,

$$\Gamma_p(\kappa) = \pi^{p(p-1)/4} \prod_{j=1}^{p} \Gamma\left(\kappa - (j-1)/2\right), \tag{23}$$

where $\kappa > (p-1)/2$ is a shape parameter, $\boldsymbol{\Sigma} \in Sym^+(p)$ is a scale parameter and $|\cdot|$ stands for the matrix determinant. If each packet of spins is represented by such a distribution, the total distribution in the voxel for N_p distinct packets is given by the weighted sum :

$$P(\boldsymbol{D}) = \sum_{j=1}^{N_p} f_j P_{\kappa_j, \boldsymbol{\Sigma}_j}, \tag{24}$$

f_j being the weight of the packet j [37]. Note that, once the different packets of spins have been isolated, it is possible to compute all the DTI metrics for each of them. Their weighted sum is then a generalisation of the classic DTI metrics, overcoming the problem of voxel homogeneity.

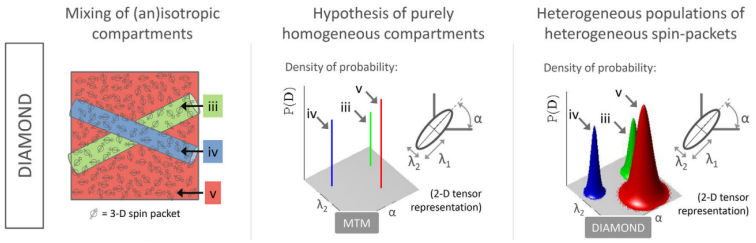

Figure 1.21: The DIAMOND spin packets and their distribution illustrated. Adapted from [37].

Benefits and limitations

Compared with the models previously discussed, DIAMOND has the advantage of proposing a model that is less onerous in terms of assumptions. This, coupled with the possibility of considering a number N_p of compartments in the voxel, makes the results of the model extremely convincing when tested in simulation. On the other hand, the number N_p must be imposed by the user, and it is common practice to set this number at 3 compartments: 2 compartments for axon bundles along two different directions and 1 compartment for the CSF [37].

1.4.4 Microstructure Fingerprinting

The aim of Microstructure Fingerprinting (MF) is to go one step further in the realistic representation of the microstructures present in each voxel. This model is also a multi-compartment model. The idea behind MF is to take the problem in reverse and find out which tissue configuration would lead to the observed diffusion signal. It is assumed that each compartment contributes linearly, as a function of its volume fraction, to the total signal perceived. Assuming that we have the signals generated by each compartment separately, all we need to do is find the linear combination of these signals that best reproduces the reality of our diffusion imaging. In practical terms, in the MF framework, the independent signals from each fascicle are obtained by Monte-Carlo simulations. A dictionary with the results from K fascicles is created, each fascicle corresponding to a previously defined population of axons and glial cells in various directions. The dictionary is intended to be relatively exhaustive. It also includes a compartment assimilated to the cerebrospinal fluid.

Based on the dictionnary, the optimal combination is found by solving a sparse optimization problem [38, 39].

Benefits and limitations

Clearly, the greatest advantage of the MF model lies in its ability to identify a very large number of independent tissue populations, which is made possible by Monte Carlo simulations. The model has also demonstrated its effectiveness by comparing its performance with that of other multi-compartment models such as NODDI or DIAMOND. However, it is acknowledged that the model still requires some improvement [29, 38, 39].

1.5 ElikoPy

ElikoPy is a Python library developed by Quentin Dessain and Mathieu Simon during their Master's thesis. The objective the library is to provide *"a complete diffusion MRI processing pipeline that reduces common sources of artifact and captures information on the brain microstructure through multiple microstructural diffusion models"* [40] (see Figure 1.22). The notable advantage of ElikoPy is that it brings together in a single package a multitude of tools needed to process and exploit DW-MRI images. In addition, the module has been developed in such a way as to be able to process large datasets simultaneously, bringing together elements with very different acquisition characteristics [40, 41].

The use of ElikoPy is essential for this work. In particular, the pipeline has been used for the general organisation of the data, the preprocessing of the DW-MRI images, the application of a white matter mask and the calculation of the various DTI, MF, NODDI and DIAMOND metrics.

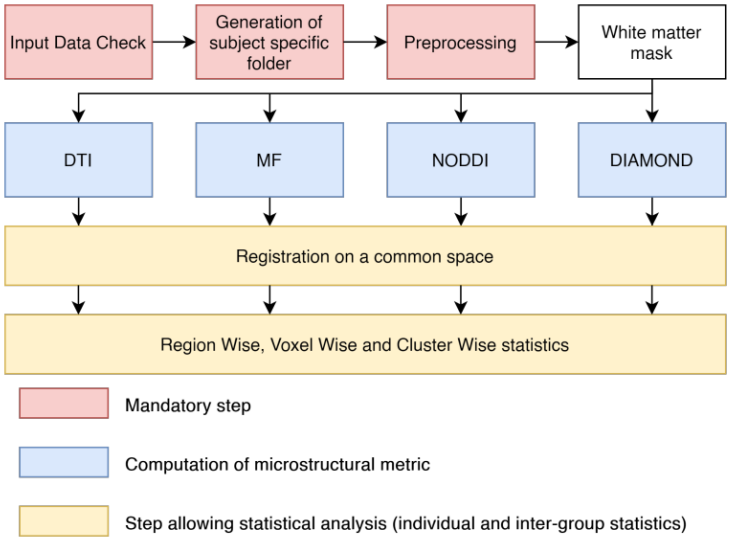

Figure 1.22: The ElikoPy pipeline's main steps. Extracted from [40].

1.5.1 Preprocessing

Before applying diffusion models to DW-MRI images, ElikoPy goes through a pre-processing stage (see Figure 1.23). The aim of this stage is to eliminate, or at least reduce, the imperfections in the images that were introduced during acquisition. Details of these artefacts can be found in the section 1.2.5 of this document.

We present here only the corrections used in this work.

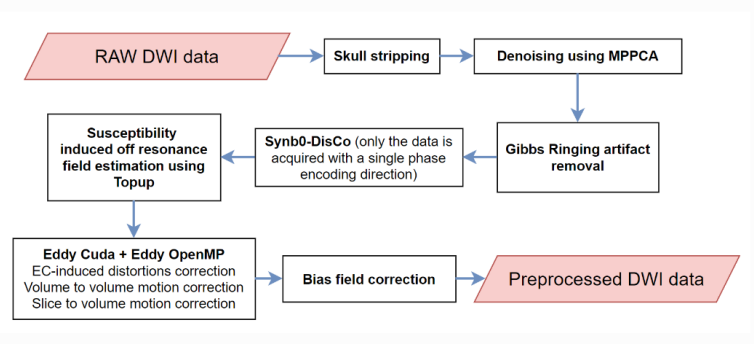

Figure 1.23: The ElikoPy preprocessing stage in details. Extracted from [41].

Skull stripping

The brain extraction from the skull and other tissues is the only mandatory step of the preprocessing stage in ElikoPy as it is required before using other image processing algorithms. Its second purpose is to increase the efficiency of subsequent processing steps.

In practical terms, extraction is achieved by applying a mask to the image (see Figure 1.24). The mask in question is computed using the function `median_otsu` from DiPy[2] [41].

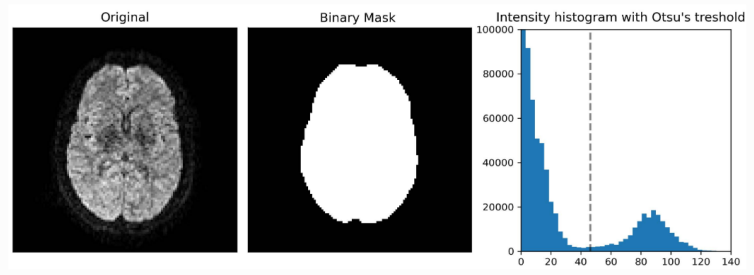

Figure 1.24: Exemple of brain extraction. Extracted from [41].

Denoising

The presence of noise is often noted in magnetic resonance images. It is generally advisable to reduce this noise in order to increase the signal-to-noise ratio of the image. Elikopy uses the Marchenko-Pastur PCA technique already implemented in DiPy to denoise the image. This choice is motivated by the speed of the algorithm and its ability not to introduce blur or additional artefacts [41].

Motion, Eddy current and susceptibility distortion correction

Motion artefacts, Eddy currents artefacts and magnetic susceptibility artefacts have different origins, as explained in section 1.2.5, but cause similar effects in the image; in this case displacements and deformations. These artefacts can therefore be corrected in the same way simultaneously. To carry out the correction, ElikoPy estimates the susceptibility distortions via

[2]DiPy beeing another Python library designed for diffusion imaging.

a generated synthetic volume based on a T1 structural image and the Eddy current distorsions using Eddy FSL which works by making a prediction about how each diffusion volume should look and comparing the prediction to the observed data [40, 41].

1.5.2 Diffusion models and metrics

Once the preprocessing stage is complete, it is possible to fit the diffusion models and calculate the corresponding metrics.

While DTI metrics are taken directly from the DiPy python package, NODDI metrics are evaluated using the Dmipy python package. The default values for NODDI parameters are set to :

$$d_{||} = 1.7 \times 10^{-9} m^2 s^{-1} \qquad \text{and} \qquad d_{iso} = 3.0 \times 10^{-9} m^2 s^{-1}.$$

DIAMOND metrics are calculated according to a code proposed by B. Scherrer, DIAMOND's main contributor. By default, ElikoPy limits the model to two neurite fascicles in addition to the CSF compartment.

The Microstructure Fingerprinting metrics use the DIAMOND tensor files to estimate the number of fascicles that will be needed. The implementation of MF script is based on the python library coded by G.Rensonnet, the main MF author. All brain tissue properties are provided by a dictionary of Monte Carlo simulation results precomputed by the author [40].

2 Method

The aim of this section is to establish the practical framework of this work. We retrace the various stages of data processing, describe the implementation choices made and establish some new concepts.

2.1 Data description

The data used in this study were collected by Professor Yves Vandermeeren and his team as part of a research project aimed at demonstrating the effectiveness of a new rehabilitation treatment. These data, kindly loaned to us, correspond to 32 patients with ischaemic stroke. For each of these patients, we have several post-stroke observations at different stages of rehabilitation. Originally, we should have had 4 observations per patient spread over time, but due to the withdrawal of some patients during the study, as well as some unusable data, we ended up with 1 to 4 observations per patient, leading to a total of 104 observations.

Each piece of data gathers a multi-shell DW-MRI image used to calculate the different diffusion metrics; a T1-weighted MRI image required for ElikoPy's pre-processing stage; and the patient's results from the 6-Minute Walk Test.

2.2 Data preprocessing

First of all, it is necessary to pre-process the DW-MRI data in order to remove potential artefacts and improve image quality. This stage is fully supported by the ElikoPy pipeline described in section 1.5.

During this preprocessing stage have been performed :

- A denoising of the images
- A correction of the motion artefacts
- A correction of the Eddy currents artefacts
- A correction of the magnetic susceptibility field artefacts based on the T1-weighted image

Because the correction of the Gibbs ringing tends to blur the images, it was chosen not to correct any Gibbs artefact as the visual inspection of the data suggested that it was not necessary. As a result, we use Elikopy's `preproc()` function with the following parameters : `denoising=True`, `forceSynb0DisCo=True` and `eddy=True`.

2.3 Microstructural metrics

Data pre-processing is directly followed by the application of the models described in section 1.4. Once again, the ElikoPy pipeline is used.

The DTI model is obtained using the `dti()` function from ElikoPy and the metric's maps used for further analyses are:

FA The Fractional Anisotropy map

MD The Mean Diffusivity map

RD The Radial Diffusivity map

AD The Axial Diffusivity map

Further information about the model can be found in section 1.4.1 and an example of the different metrics is available in Figure 2.1.

The NODDI model and its metrics are obtained using the `noddi()` function from ElikoPy. Metrics used for further analyses are:

ODI The fiber Orientation Dispersion Index map

fextra The extra cellular volume fraction map

fintra The intra cellular volume fraction map

fiso The Free water volume fraction map

fbundle The Fiber bundles volume fraction

Section 1.4.2 provides more information on this model and an example of the results is shown in Figure 2.3.

The DIAMOND model and its associated metrics are provided by the `diamond()` function of the ElikoPy pipeline. The selected features of the DIAMOND model are:

f0 The first fiber population volume fraction map

f1 The second fiber population volume fraction map

csf The cerebrospinal fluid volume fraction map

t0 The first fiber population tensors map

t1 The second fiber population tensors map

All needed information about DIAMOND can be found in section 1.4.3. More than the volume fraction maps, the major interest of this model resides in the fact that it allows the computation of enhanced DTI metrics. Indeed, since DIAMOND is able to handle and distinguish two fiber populations in one single voxel, it can be used to partially overcome the limitations of DTI in case of crossing fibers. This is done computing the weighted mean of DTI metrics for both fiber populations as described in equation 25.

$$wM = \frac{f0 \times pM\,(t0) + f1 \times pM\,(t1)}{f0 + f1}, \tag{25}$$

where wM stands for "weighted Metric" (the "DIAMOND-enhanced"[3] DTI metric) and pM for "population Metric" (the DTI metric computed for the considered fibre population). We thus obtain the following features :

diamond_FA The DIAMOND-enhanced Fractional Anisotropy map

diamond_MD The DIAMOND-enhanced Mean Diffusivity map

diamond_RD The DIAMOND-enhanced Radial Diffusivity map

diamond_AD The DIAMOND-enhanced Axial Diffusivity map

An example is displayed in Figure 2.2.

The last model used is the Microstructure Fingerprinting model (MF). It is obtained via the `fingerprinting()` function from ElikoPy. Please note that, in the case of this study, we forced the estimation of the fascicles directions using the `odf_csd` function from Elikopy. The metric's maps used for further analyses are:

DIFF_ex_tot The total extra-axonal diffusivity map

frac_csf The cerebrospinal fluid volume fraction map

fvf_tot The fiber volume fraction map of all fascicles

Further information about the model can be found in section 1.4.4 and an example of the different metrics is available in Figure 2.4.

[3]This name does not actually reflect the truth as DIAMOND is model completely different from DTI

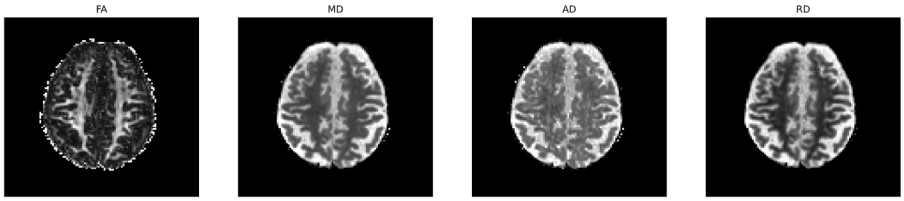

Figure 2.1: DTI metrics maps (`FA - MD - AD - RD`)

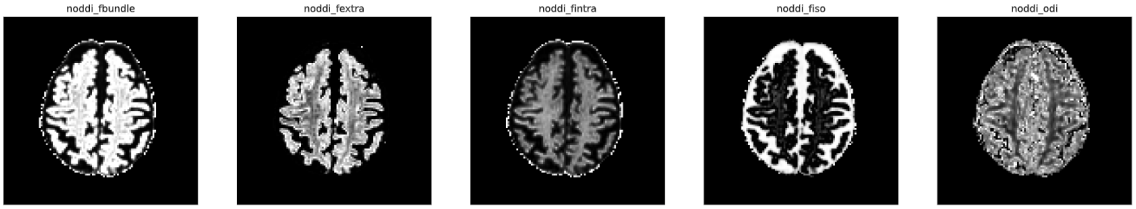

Figure 2.2: NODDI metrics maps (`fbundle - fextra - fintra - fiso - ODI`)

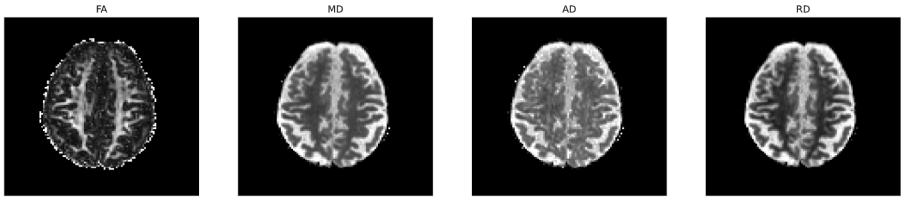

Figure 2.3: DIAMOND-enhanced DTI metrics maps (`diamond_FA - diamond_MD - diamond_AD - diamond_RD`)

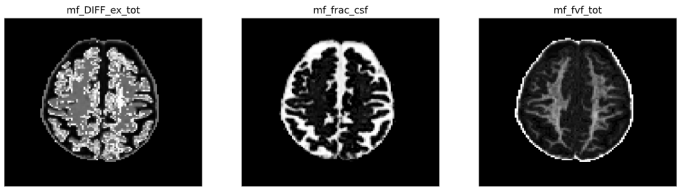

Figure 2.4: MF metrics maps (`DIFF_ex_tot - frac_csf - fvf_tot`)

2.4 Data registration and regions of interest

Based on the previously calculated metric maps, a framework needs to be established to enable the elements in the dataset to be compared.

2.4.1 ROI selection

It goes without saying that studying the metrics for the whole brain of each patient would be of little use. It is more appropriate to divide our data into regions of interest sharing anatomical properties.

For the purposes of this study, it would have been too ambitious to spread out across different brain regions. That is why, although it would have been interesting to extend our study to different brain areas, it was decided to restrict the analysis of diffusion metrics to the corticospinal tract alone. The corticospinal tract (CST) is composed of white matter. It originates in the fronto-parietal cortices (including the primary motor cortex, secondary motor area and somatosensory cortex) and connects them to the spinal cord. It is considered to be the main neural pathway for voluntary motor functions, making it a region of strategic importance for this study [42].

2.4.2 Atlas

In our data, regions of interest (ROI) are identified using atlases available on the FSL website [43]. More specifically, we use here the *HCP1065 standard-space DTI templates* and its FA map as a support for image registration as well as the *XTRACT HCP Probabilistic Tract Atlases* for the identification of the left and right corticospinal tracts.

The masks of the *XTRACT HCP Probabilistic Tract Atlases* are probabilistic, meaning that a value between 0 and 100 is associated to each voxel of the mask. This value represents the probability in % that the voxel belongs to the designated tract. In this work, it has been decided not to apply a threshold to these masks, but to use their probabilities as weights in the calculation of metrics on the considered region. The weighted mean of a metric on a region is thus given by :

$$M_R = \frac{\sum\limits_{i\in\Omega} p_R^i \times m^i}{\sum\limits_{i\in\Omega} p_R^i}, \tag{26}$$

where M_R is the averaged metric on region R, Ω the whole domain of the image, p_R^i the value of the mask of the region R on the voxel i and m^i the value of the metric on the voxel i.

2.4.3 Registration

As mentioned above, our process involves a registration step in which the atlases are projected from the reference space into the patient's space. For each patient, the statistical analysis of the metrics is therefore carried out in his or her own reference space. Registration aims to compute the projection of a DTI FA map provided by the atlases onto the patient's DTI FA map (see

Figure 2.5). The transformation is then applied to the desired masks of the ROI's.

The transformation is computed using the DIPY python module and its implemented functions. A pipeline is used to compute different transformations one after the other. This allows finer results for the registration. The program respectively computes a translation to align the "centers of mass" of the two images, a translation maximizing the mutual information between the images, a rigid transformation (rotation and translation) maximizing the same metric and an affine registration. At last, a Symmetric Diffeomorphic Registration optimizing the MI between the two references is used to reach the final transformation.

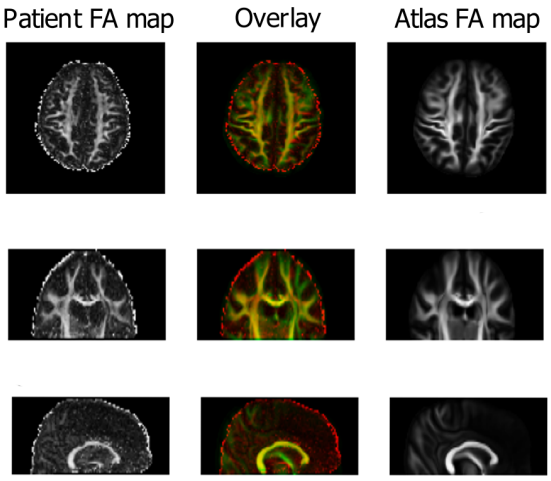

Figure 2.5: Example of a projection of the Atlas FA map onto a patient's FA map.

Because the acquisition parameters were not the same for all the patients of the study, parts of the brain, in some cases, have been truncated in the images. For the sake of comparability, it was necessary to compute the future metrics of diffusion in a common area for all patients. That is why a mask has been applied on all images to only keep the parts of the brain that have been correctly acquired for every patients as shown in Figure 2.6.

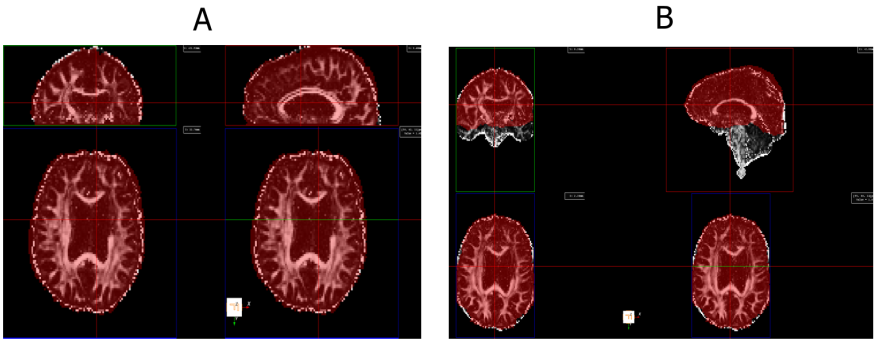

Figure 2.6: FA maps of two different patients (A and B). The brain has been truncated in patient A. A mask (in red) of the common area is added to the brains.

2.5 Metrics analysis and measures definition

The maps generated by the diffusion models assign at least one value per voxel in an image. Even if we limit ourselves to voxels located in a region of interest, we still retain a few hundred values per metric. The key is then to condense all this information into fewer features to allow a better and easier interpretation of results and to use statistical models to highlight trends between dMRI data and the brain's health of the patient in the specified area. To do this, we need to define a series of measures that summarise maps of distribution metrics.

2.5.1 Traditional metrics analyses

A common measure for representing the values of a metric over an entire ROI is its mean value. Several studies have already shown a form of dependency between the mean value of diffusion metrics and the results obtained by a patient in a clinical test after a stroke [44]. Generally, these mean values are calculated separately for each side of the brain. The values obtained for the ipsilesional and contralesional sides are then compared by taking the difference or ratio.

Giving M a metric averaged on a region, we define two measures borrowed from [29] to compare ipsilesional and contralesional sides: rM the metric ratio given by equation 27 and aM the metric asymmetry defined by equation 28.

$$rM = \frac{M_{ipsilesional}}{M_{contralesional}}. \tag{27}$$

$$aM = \frac{M_{ipsilesional} - M_{contralesional}}{M_{ipsilesional} + M_{contralesional}}. \tag{28}$$

2.5.2 Metrics distribution analyses

In this work, we want to try a different approach. Although the use of average values has proved effective, we are trying to define new complementary and/or alternative measures to summarise the diffusion metrics over a region of interest. These new methods are based on the construction and analysis of histograms associated with the metrics over the area under consideration. We call this new method the *Metric distribution analysis* (MDA)

Histogram calculation

Considering a metric M over a region R (here the corticospinal tract), we want to construct a histogram of M showing the proportions in which the different values of M are represented in the zone under consideration.

Note that before calculating the histogram, we need to remove any outliers from M over R. After inspection, we can see that over an entire region, it is common that only a few voxels have truly extravagant values. In order to limit data alteration, we therefore chose to truncate the values of M at its 99^{th} percentile. As a result, any value greater than $M_R^{99\%}$ will be reduced to

$M_R^{99\%}$. The histogram will therefore be defined on $[M_R^{min}, M_R^{99\%}]$.

This interval is divided into bins. For the purposes of this study, one hundred bins were chosen to divide the interval as a compromise between the number of bins that needed to remain optimally small and the continuous representation of the metric distribution. B^i the i$^{\text{th}}$ bin of the distribution is defined as the interval $[M_R^{min} + i \times E_{bins}, M_R^{min} + (i+1) \times E_{bins}]$ where E_{bins} is the width of the bin (the range of the bin's interval) and is equal to $\frac{M_R^{99\%} - M_R^{min}}{100}$. Again, the probabilities given by the atlas maps are used as weights to compute the density of the distribution lying in the bins. Hence, the value V^i of the density lying in the bin B_i is given by :

$$V_R^i = \frac{\sum\limits_{j|M^j \in B_R^i} p_R^j}{\sum\limits_{j \in \Omega} p_R^j}, \tag{29}$$

with Ω the whole domain of the image, p_R^j the value of the mask of the region R on the voxel j and M^j the value of the metric on the voxel j. The result of this procedure can be represented as an histogram or a distribution function as shown in Figure 2.7. From this reorganised data, it is possible to define new quantities that will measure the divergence between the ipsilesional and contralesional distributions.

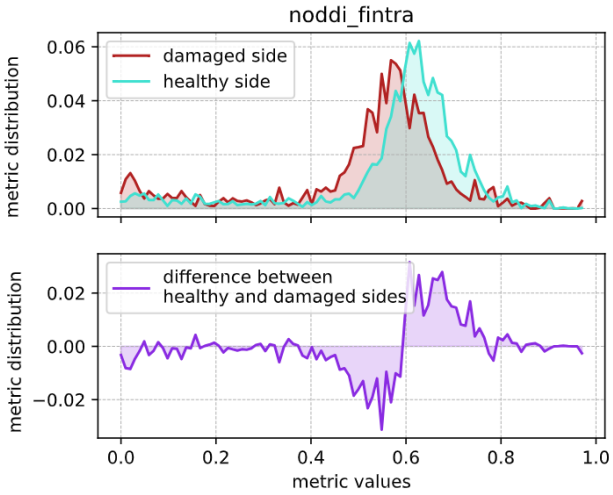

Figure 2.7: Estimated distribution functions for the `noddi_fintra` metric of a patient over the corticospinal tract on its damaged and healthy sides along with their difference.

Jensen-Shannon distance

The Jensen-Shannon distance is the square root of the Jensen-Shannon divergence. This divergence is, in probability theory and statistics, a measure of the similarity between two probability distributions [45]. The distance between the probability distributions X and Y is given by:

$$JSD(X,Y) = \sqrt{\frac{D(X||Z) + D(Y||Z)}{2}}, \tag{30}$$

where Z is is the pointwise mean of X and Y and D is the Kullback-Leibler divergence. We use the *scipy* Python module and its function `jensenshannon()` to compute it for the ipsilesional and contralesional distributions [46].

Dynamic Time Warping distance

Usually used to measure similarity between two temporal sequences, here the Dynamic Time Warping (DTW) distance is used as a measure of similarity between the ipsilesional and contralesional distributions. The DTW tries to find the best global alignment of the two distributions exploiting temporal (along the x-axis) distortions between them [47]. The distance is computed using the `dtw()` function the *dtaidistance* package in Python [48].

Mean Curve Distance

The Mean Curve Distance (MCD) is the most direct way to define a distance between the two distributions. It is given by the mean of the pointwise difference between the contralesional and the ispsilesional distributions taken to the absolute (see Figure 2.8), which gives :

$$MCD = \frac{1}{100}\sum_{i=1}^{100} |V_i^{contralesional} - V_i^{ipsilesional}|. \tag{31}$$

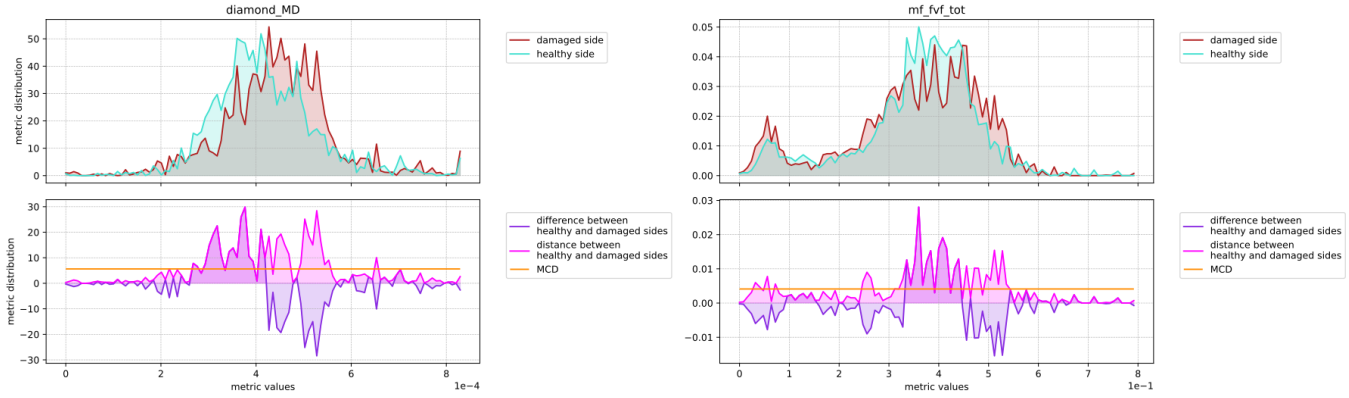

Figure 2.8: Representation of MCD on distribution graphs.

Standard deviation of the Curve Distance

The Standard deviation of the Curve Distance (SCD) (see Figure 2.9) attempts to identify whether the differences between the two distribution curves are evenly distributed along these curves or whether, on the contrary, they are condensed at certain points of great divergence. Its formula is given by :

$$SCD = \sqrt{\frac{1}{100}\sum_{i=1}^{100}\left(\left|V_i^{contralesional} - V_i^{ipsilesional}\right| - MCD\right)^2}. \tag{32}$$

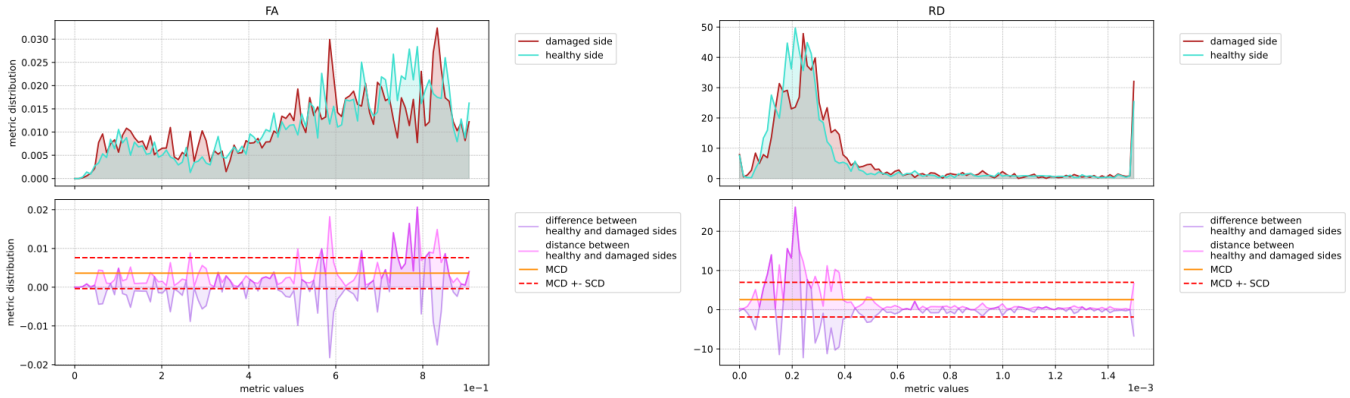

Figure 2.9: Representation of SCD on distribution graphs.

Cumulative Curve Difference

If the histograms can be assimilated to distribution functions, the Cumulative Curve Difference (CCD) is the integral over the domain of the difference between the contralesional and the ipsilesional cumulative distributions of the metric (see Figure 2.10). Its aim is to show the presence of shifts between the distribution functions. The measure is computed via the formula :

$$CCD = \sum_{i=1}^{100} \left(\sum_{j=1}^{i} V_j^{ipsilesional} - V_j^{contralesional} \right). \tag{33}$$

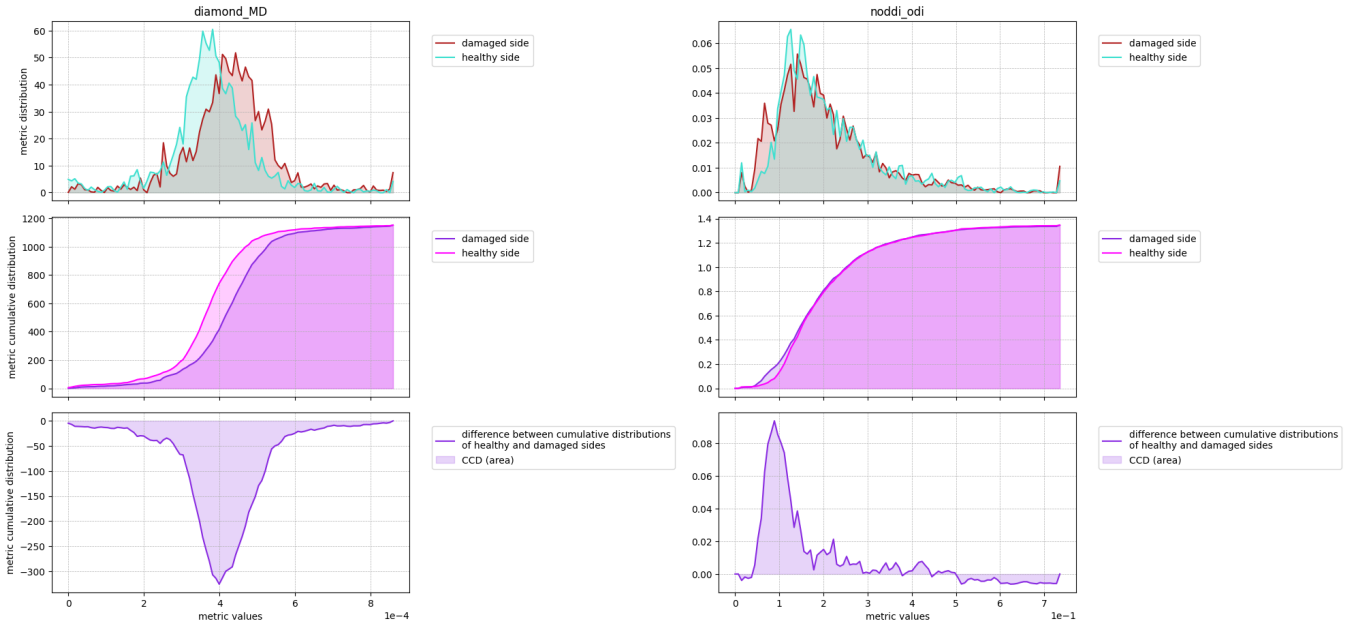

Figure 2.10: Representation of CCD on distribution graphs.

2.6 Statistics

The aim of this study is to determine whether the use of new MDA measures provides any additional information about a patient's state of health based on their diffusion metric maps. To do this, we use patients' results in the *6-Minute Walk Test* as a "target" variable. We consider that this variable represents the patient's general state of health (in relation to the stroke).

The first step in our analysis is to confirm the relevance of the various measures in relation to the target. On this basis, based on our data, for each measure applied to each metric, we use a statistical test to demonstrate a linear dependency between the considered measure and the target. Because the relationship between the results in the *6-Minute Walk Test* and the metrics could be non-linear, we also decide to compute the mutual information the measures and the target. In addition, in order to determine whether our newly introduced measures provide access to additional information to that already provided by the traditional measures (rM and aM of Equations 27 and 28), we also correlation coefficient showing a linear relationship between them.

The statistical tests performed are F-tests applied to the null hypothesis: "Measure m of metric M is not linearly dependent on target". The test is said to be significant if it shows a p-value of less than 0.05. In this case, the null hypothesis can be rejected, ensuring the existence of a linear dependency. The F-scores of the tests and their associated p-values are obtained using the `f_regression()` function of the Python package *scikit-learn* [49]. The correlation coefficient between measures is computed by the function `r_regression()` of the Python package *scikit-learn*[49].

The mutual information (MI) between two variables X and Y quantifies the consequences of a loss of uncertainty in one of the variables on the uncertainty of the other variable. Since the uncertainties of two (linearly or non-linearly) dependent variables are strongly related, their mutual information is large. Conversely, two independent variables have, in principle, zero MI [50]. The mutual information between the target variable and the measures are computed using the `mutual_info_regression()` function of *scikit-learn* [49].

Note that, before running the analysis, outliers are removed from the data and features are centered on their mean and normalized by their standard deviation. A sample S of the variable X is categorized as an outlier if :

$$S < X_{25\%} - 1.5 \times (X_{75\%} - X_{25\%}) \qquad \text{or} \qquad S > X_{75\%} + 1.5 \times (X_{75\%} - X_{25\%}). \tag{34}$$

3 Results

This section presents the results and statistics of the MDA's performance in assessing patient health on the basis of diffusion metric maps over the corticospinal tract. These results were obtained using the method described in Section 2 and will be discussed in Section 4. The results are grouped here by model and diffusion metric. For more information, the measure-target scatter plots of each measure for each metric are available in Appendix A, B, C and D for the DTI, DIAMOND, MF and NODDI models respectively.

3.1 DTI results

3.1.1 Fractional Anisotropy

The results shown in Figure 3.1 indicate that there is indeed a linear dependency between the target and the SCD, MCD, DTW and *Jensen-Shannon* measurements. Significance is established for a p-value of less than 0.05. On the other hand, we note that the mutual information between these measures and the results of the *6-minute walk Test* is lower than that of the CCD, rM and aM measures. This suggests the possibility of a non-linear relationship between these measurements and the target.

Concerning the linear dependence of the MDA measurements and those of rM and aM, Figure 3.2 shows a relative correlation between the measurements. Indeed, it indicates a coefficient greater than 0.7 (in absolute value) for all the MDA measurements with the exception of DTW and SCD.

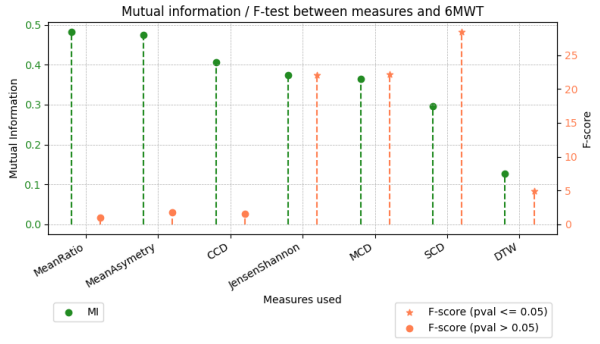

Figure 3.1: Mutual information and F-test results of the measures for the FA map on the corticospinal tract.

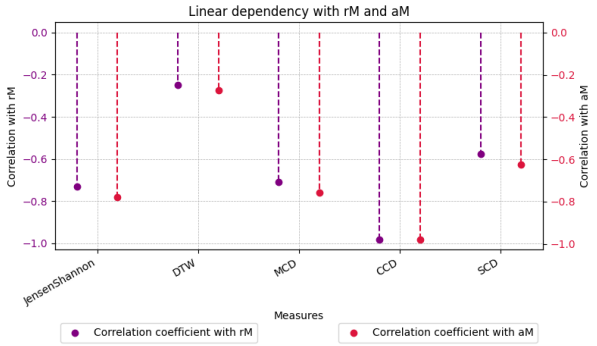

Figure 3.2: Correlations between MDA measures and traditionnal measures for the FA map on the corticospinal tract.

3.1.2 Axial Diffusivity

Based on Fig. 3.3, we can see that the CCD, rM and aM measurements appear to be linearly related to the target. The MI tends to confirm these results.

Figure 3.4 shows a strong correlation between the CCD, and aM and rM. This correlation still seems to be present but is weaker for the other measures.

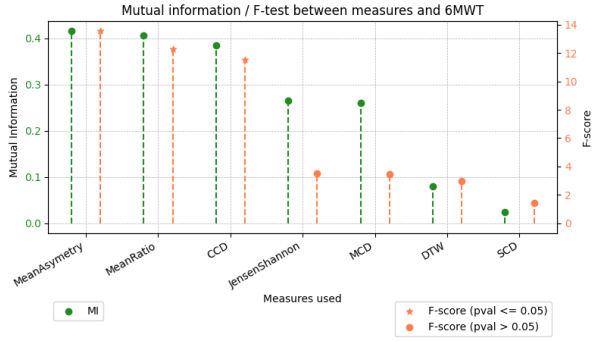

Figure 3.3: Mutual information and F-test results of the measures for the AD map on the corticospinal tract.

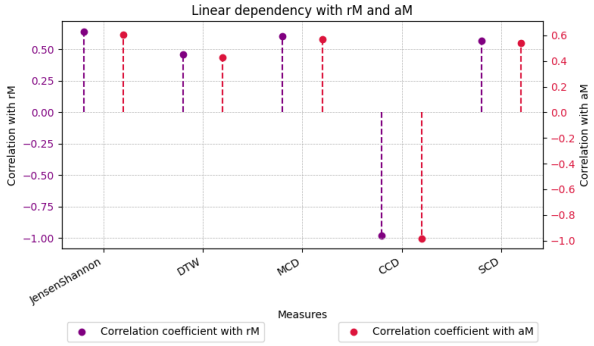

Figure 3.4: Correlations between MDA measures and traditionnal measures for the AD map on the corticospinal tract.

3.1.3 Radial Diffusivity

Figure 3.5 shows a linear dependence between the target and the SCD, MCD, DTW and *Jensen-Shannon* measures, with a demarcation for the SCD. These same measures also share the greatest mutual information with the target, confirming the dependency. We notice that the MI for the aM and the rM are relatively close from the MI of the DTW which might mean a non-linear relationship with the target.

In Figure 3.6, we can see a relatively strong correlation between all the measurements and the aM and rM, the least marked correlation being that of the SCD with coefficients close to 0.5.

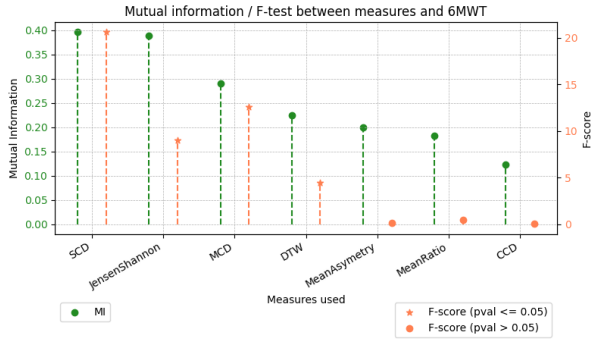

Figure 3.5: Mutual information and F-test results of the measures for the RD map on the corticospinal tract.

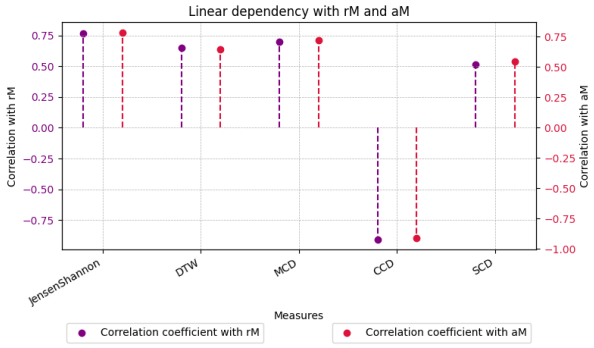

Figure 3.6: Correlations between MDA measures and traditionnal measures for the RD map on the corticospinal tract.

3.1.4 Mean Diffusivity

Looking at Figure 3.3 we can see that only the DTW and the MCD are linearly dependent on the results of the 6-Minute Walk test. Moreover, this correlation does not seem very strong for the MCD in view of its F-score. Nevertheless, these two measures are also those with the highest MI values. They are closely followed by the *Jensen-Shannon* which might be non-linearly

related to the target.

The Figure 3.8 shows a relative correlation between the *Jensen-Shannon distance*, the MCD, the SCD and the aM and rM measures. The correlation coefficient however does not exceed 0.6.

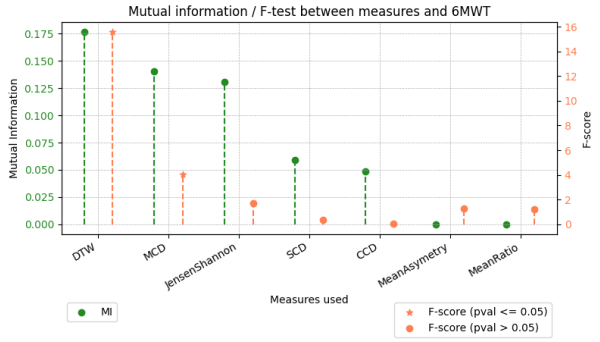

Figure 3.7: Mutual information and F-test results of the measures for the MD map on the corticospinal tract.

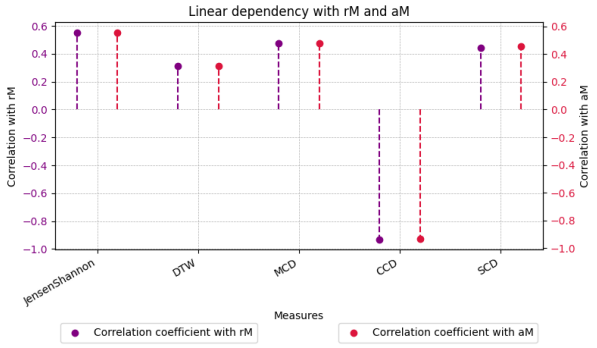

Figure 3.8: Correlations between MDA measures and traditionnal measures for the MD map on the corticospinal tract.

3.2 DIAMOND Results

3.2.1 DIAMOND-enhanced Fractional Anisotropy

Figure 3.9 shows a verified linear dependence between the target and the *Jensen-Shannon*, MCD, DTW and SCD measurements. That being said, given the F-score values, this dependency does not appear to be very strong. It is also interesting to note that all other measures have a higher MI than the correlated variables, indicating a possible non-linear connection or the weakness of the linear dependence of the previously cited features.

Concerning the linear dependence of the MDA measurements and those of rM and aM, Figure 3.10 clearly shows that all MDA measures seems to be correlated with the rM and the aM (correlation coefficient superior to 0.7).

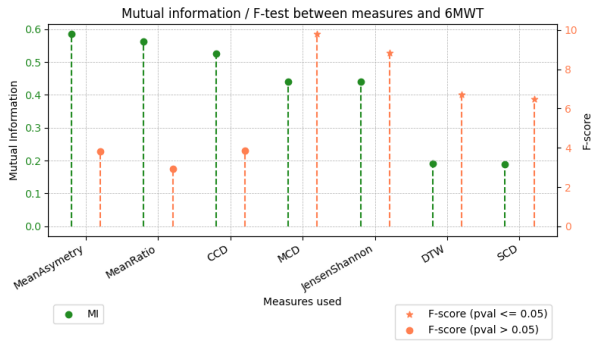

Figure 3.9: Mutual information and F-test results of the measures for the DIAMOND-enhanced FA map on the corticospinal tract.

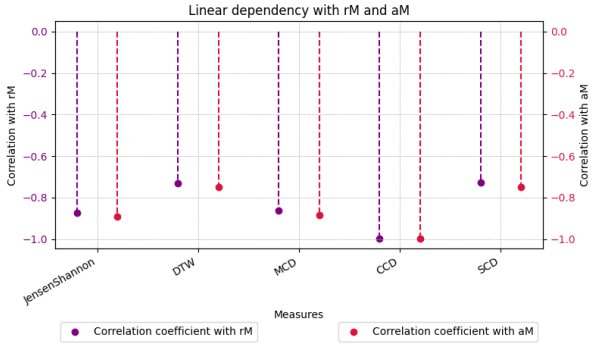

Figure 3.10: Correlations between MDA measures and traditionnal measures for the DIAMOND-enhanced FA map on the corticospinal tract.

3.2.2 DIAMOND-enhanced Axial Diffusivity

Based on Fig. 3.11, we can see that all measures except the MCD appear to be linearly related to the target. The MI tends to confirm these results and shows a gap between the 3 first features and the others.

Figure 3.12 still shows a strong correlation between the CCD and the aM and rM measures but the other measures have very weak correlation coefficients with the aM and the rM lying between 0.1 and 0.3.

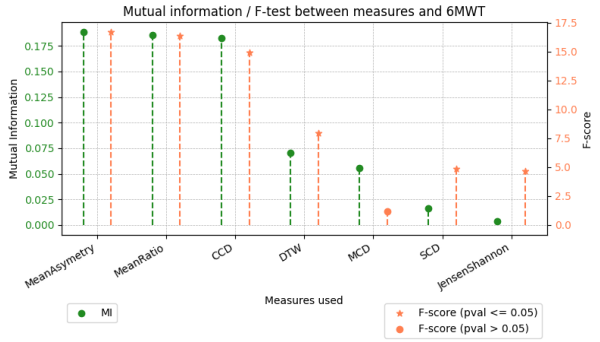

Figure 3.11: Mutual information and F-test results of the measures for the DIAMOND-enhanced AD map on the corticospinal tract.

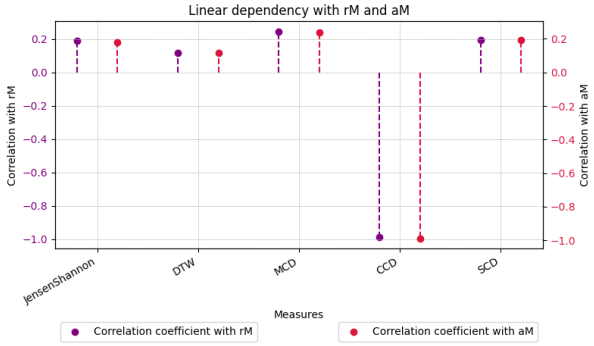

Figure 3.12: Correlations between MDA measures and traditionnal measures for the DIAMOND-enhanced AD map on the corticospinal tract.

3.2.3 DIAMOND-enhanced Radial Diffusivity

Looking at Figure 3.13, we can see that the features of aM, rM and MCD appear to be linearly related to the target. Additionally, the F-score seems to indicate that this link is relatively weak. As the MI values are relatively uniform, it is also difficult to tell whether a measure has a non-linear relationship with the target.

As for the correlation between the MDA measurements and those of aM and rM, Figure 3.14 seems to indicate that there is a relatively strong correlation between all the features and the usual measures, the CCD having once again the strongest correlation.

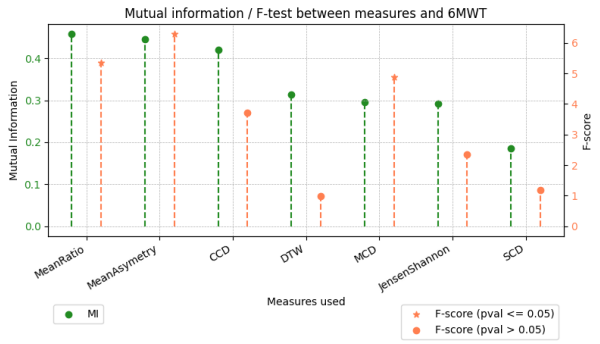

Figure 3.13: Mutual information and F-test results of the measures for the DIAMOND-enhanced RD map on the corticospinal tract.

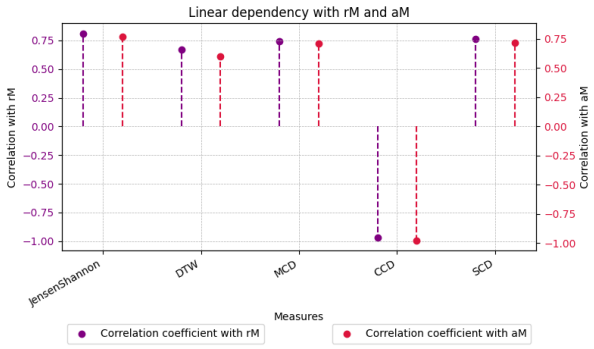

Figure 3.14: Correlations between MDA measures and traditionnal measures for the DIAMOND-enhanced RD map on the corticospinal tract.

3.2.4 DIAMOND-enhanced Mean Diffusivity

Analysis of Figure 3.15 reveals a linear relationship between all measurements with the exception of the SCD. Although, considering the F-scores, it seems that most of the features have a weak

linear bond with the target. The MI tends to say that it may exist a non-linear relationship between the aM, the rM, the DTW, the CCD and the target.

Analysis of the correlations in Figure 3.16 reveals a strong correlation between the CCD, the aM and the rM, as well as a more modest correlation with the *Jensen-Shannon distance*. The other measurements appear to have little or no correlation with the aM and rM.

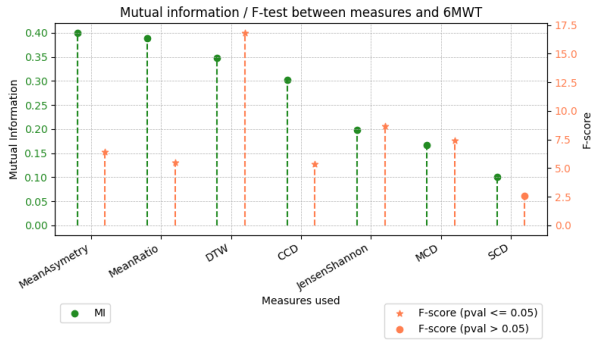

Figure 3.15: Mutual information and F-test results of the measures for the DIAMOND-enhanced MD map on the corticospinal tract.

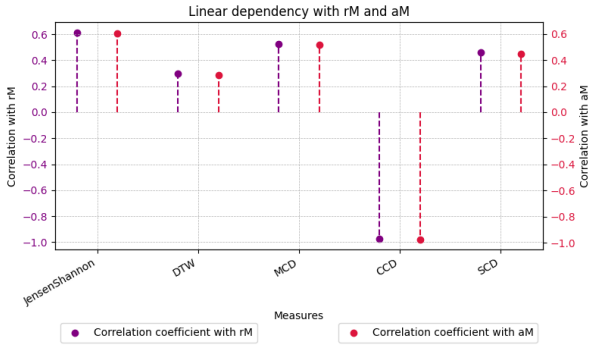

Figure 3.16: Correlations between MDA measures and traditionnal measures for the DIAMOND-enhanced MD map on the corticospinal tract.

3.3 NODDI Results

3.3.1 Intracellular volume fraction

The F-test in Figure 3.17 does not confirm any linear dependency between the features and the target. The relatively high MI values for the CCD in particular may indicate a non-linear relationship with the target, but it is impossible to state this with certainty.

Figure 3.18 shows a strong correlation between the CCD, the MCD, the *Jensen-Shannon distance*, the aM and the rM, and weaker correlations for the other features.

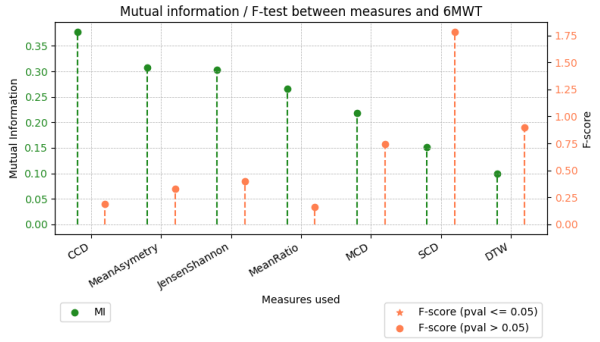

Figure 3.17: Mutual information and F-test results of the measures for the `fintra` map on the corticospinal tract.

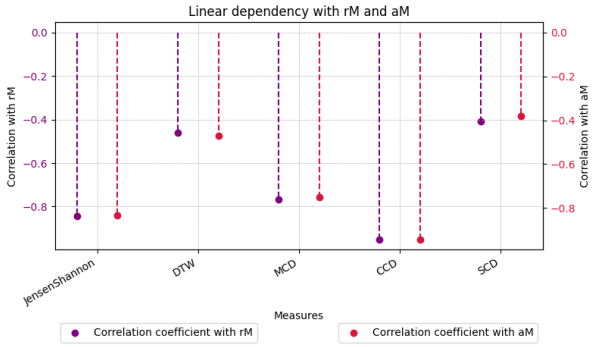

Figure 3.18: Correlations between MDA measures and traditionnal measures for the `fintra` map on the corticospinal tract.

3.3.2 Extracellular volume fraction

The results in Figure 3.19 show a linear relationship between the CCD, the aM, the rM, the DTW and the SCD with the target. On the other hand, the values of mutual information are almost as high for the linearly independent measures as for those mentioned above. This could potentially indicate a non-linear relationship between those features and the target. However, the visual inspection of Figure D.2 does not suggest this kind of link with the target.

From Figure 3.20, we can apparently only draw a correlation between CCD and traditional measurements.

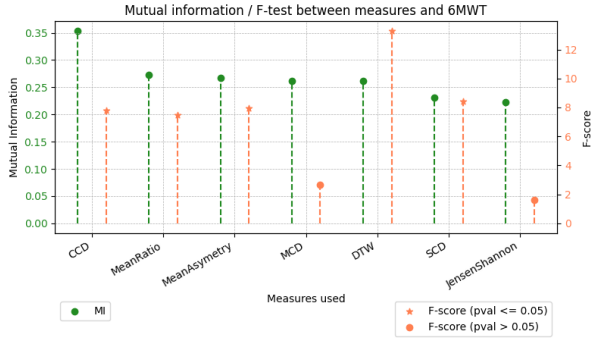

Figure 3.19: Mutual information and F-test results of the measures for the `fextra` map on the corticospinal tract.

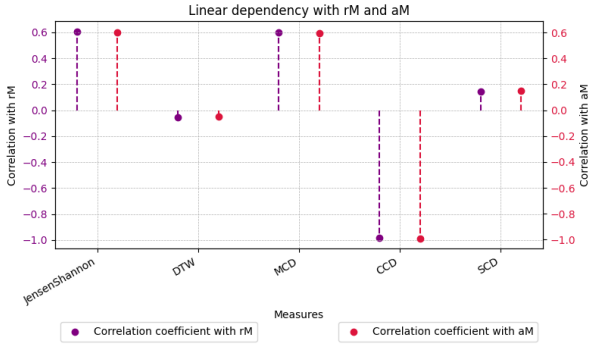

Figure 3.20: Correlations between MDA measures and traditionnal measures for the `fextra` map on the corticospinal tract.

3.3.3 Free water volume fraction

The results of the F-test in Figure 3.21 indicate no linear dependence between the features and the target. Given the MI values of all features, it would seem fairly safe to claim that there is no link between them and the target either.

Excepted for the *CCD* Figure 3.22 tends to show weak or totally non-existent correlations between MDA measurements and traditional measurements.

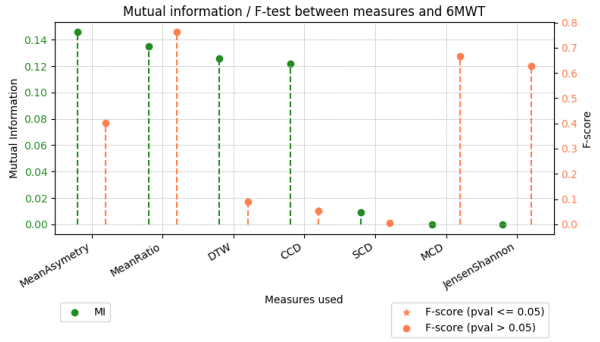

Figure 3.21: Mutual information and F-test results of the measures for the `fiso` map on the corticospinal tract.

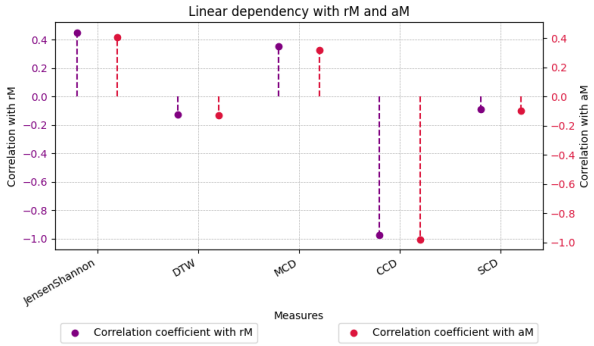

Figure 3.22: Correlations between MDA measures and traditionnal measures for the `fintra` map on the corticospinal tract.

3.3.4 Fiber bundles volume fraction

The results obtained by analysing the measurements for the fiber bundles volume fraction and presented in Figures 3.23 and 3.24 are very similar to those obtained for the free water volume fraction.

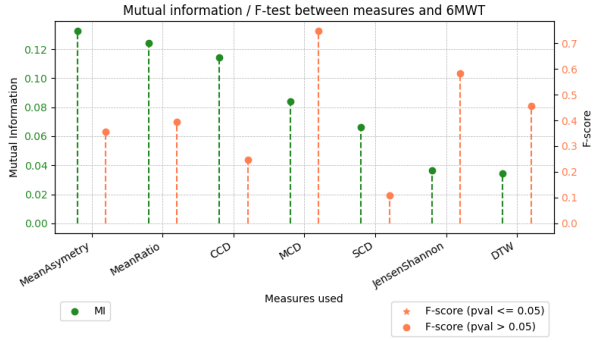

Figure 3.23: Mutual information and F-test results of the measures for the `fbundle` map on the corticospinal tract.

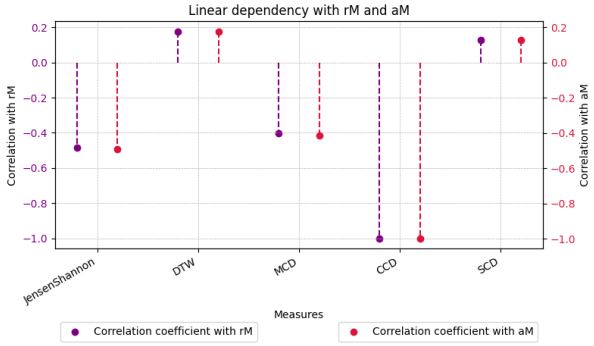

Figure 3.24: Correlations between MDA measures and traditionnal measures for the `fintra` map on the corticospinal tract.

3.3.5 Fiber orientation dispersion index

The Analysis of the measurements for the `ODI` map shows in Figure 3.25 a linear dependence of the CCD, the aM, the rM, the DTW and, to a lesser extent, the MCD on the target. Significant values for the mutual information reinforce this analysis in the case of the first three measures. Given that MI is the weakest for DTW, which appears to be strongly linearly linked to the target, it is plausible to think that there is at least a non-linear link with the target for all the measures. The visual inspection of Figure D.5 strengthens this hypothesis.

The results in Figure 3.22 show a very weak correlation for the *Jensen-Shannon distance* and the SCD, a relatively weak bond for the MCD and the DTW and again a strong correlation for the CCD.

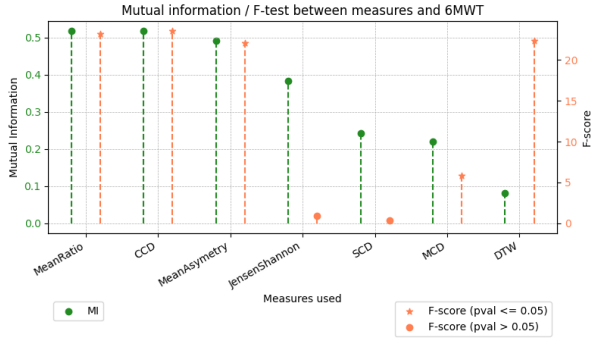

Figure 3.25: Mutual information and F-test results of the measures for the `ODI` map on the corticospinal tract.

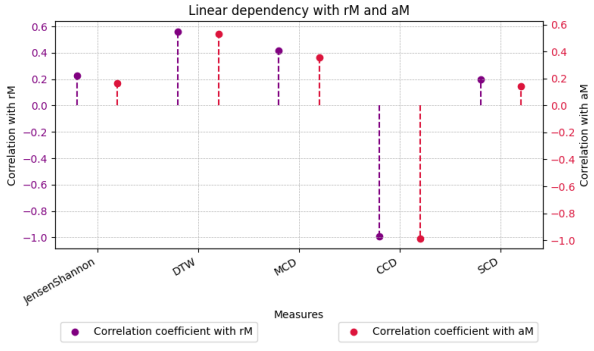

Figure 3.26: Correlations between MDA measures and traditionnal measures for the `ODI` map on the corticospinal tract.

3.4 Microstrure fingerprinting Results

3.4.1 Total extra-axonal diffusivity

The results presented in Figure 3.27 do not indicate a linear dependency between the measurements and the target. As the mutual information for the different features is relatively small, it is also difficult to derive any information about a non-linear dependency.

Figure 3.28 shows only weak correlations between the MDA and more traditional measurements. Only the *CCD* still stands out.

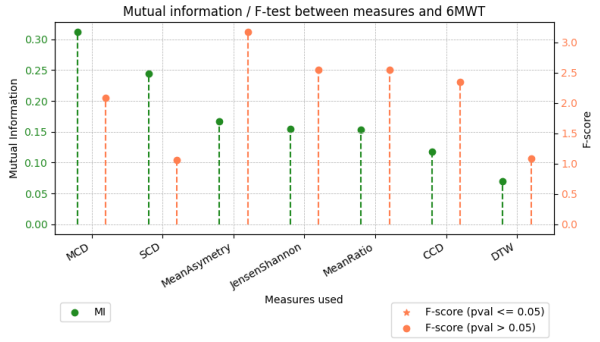

Figure 3.27: Mutual information and F-test results of the measures for the `DIFF_ex_tot` map on the corticospinal tract.

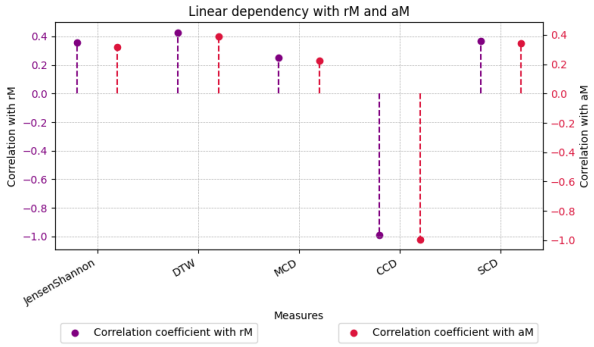

Figure 3.28: Correlations between MDA measures and traditionnal measures for the `DIFF_ex_tot` map on the corticospinal tract.

3.4.2 Cerebrospinal fluid volume fraction

Analysis of Figure 3.29 shows a linear dependence between the *CCD*, the *DTW*, the *Jensen-Shannon distance* and the target. This dependence seems to be fairly weak according to the F-scores. The features previously mentioned being the ones with the lowest MI, either all the measures have at least a weak non-linear relationship with the target, or our features are really poorly correlated with it. Visual analysis of the scatter-plot in Figure C.2 tends to support the first hypothesis.

The Figure 3.30 reflects at least a slight correlation between all the MDA measures and the traditional measures.

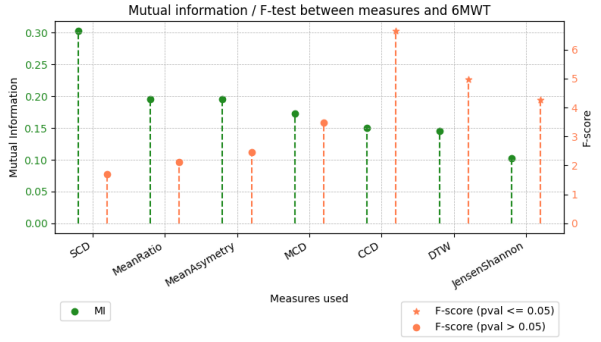

Figure 3.29: Mutual information and F-test results of the measures for the `frac_csf` map on the corticospinal tract.

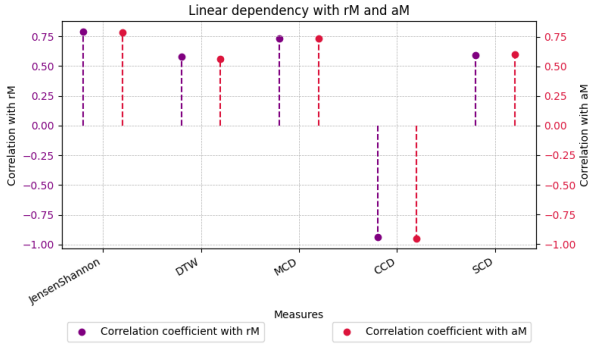

Figure 3.30: Correlations between MDA measures and traditionnal measures for the `frac_csf` map on the corticospinal tract.

3.4.3 Fiber volume fraction

According to Figure 3.31, the statistical test and the mutual information do not suggest that a feature has an established dependency with the target.

Figure 3.32 shows a strong correlation between all the newly introduced measures and the usual measures, the only exception being the *DTW*.

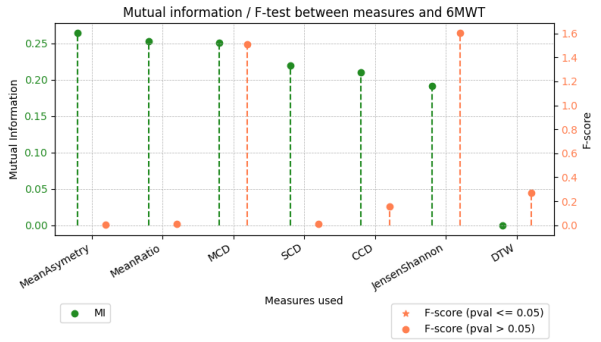

Figure 3.31: Mutual information and F-test results of the measures for the `fvf_tot` map on the corticospinal tract.

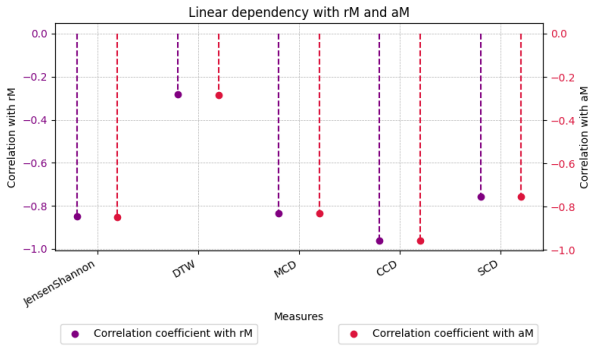

Figure 3.32: Correlations between MDA measures and traditionnal measures for the `fvf_tot` map on the corticospinal tract.

3.5 Summary

To sum up, the results are highly dependent on the metrics represented. The MF and NODDI metrics show less significance in representing the *6-minute walk test* results while DIAMOND and DTI metrics appear to be relatively closely linked to the clinical test results. Aside from that, it seems to be quite common for the MDA measures and those of aM and rM not to be simultaneously relevant (according to the statistical tests) for representing a metric. In all cases, it does not seem that one method provides unquestionably better results than the other. Sometimes, however, a strong dependency between the target and a metric is identified when the other features show a very tenuous link with the target. In addition, various correlations can be observed between the MDA measures and those based on the mean of the distributions. Even if this does happen, it seems rare for an MDA measure to be perfectly decorrelated with the aM and rM while being significant in the representation of the metric. Most often, MDA measures display correlation coefficients between 0.5 and 0.9.

4 Discussion

The results presented in Section 3 provide us with a wealth of information useful in determining whether or not MDA-based measures are relevant in the analysis of brain microstructures after ischaemic stroke. We have obtained very different results depending on the diffusion models and the metrics employed. The conclusions to be drawn from these findings are not necessarily straightforward and require in-depth investigation. In this section, we discuss and interpret these results, point out the advantages and limitations of our method and open the door to future perspectives.

4.1 Relevance of the MDA

The statistical significance of the MDA measures was found to be highly dependent on the diffusion metrics they represented. In the same way, we observe that the redundancy of the information provided by our new measures and measures such as the *mean asymmetry* and the *mean ratio* also depends on the metrics analysed. It is thus necessary to discuss these results independently with regard to the different diffusion models. We can therefore also attempt to interpret the MDA measurements and their relationship with the results of the *6-minute walk test*. The only exception is the case of the CDD which is systematically highly correlated to the aM and the rM. This phenomenon is not surprising since these measures have the same purpose, namely to demonstrate an overall shift in the values of the ipsilesional distribution compared with those of the contralesional one. It is therefore logical that these features should carry the same information.

4.1.1 The MDA for the DTI metrics

DTI metrics show great promise in favour of MDA measures. While *axial diffusivity* offers more questionable results, *fractional anisotropy, mean diffusivity* and *radial diffusivity* show the potential of certain of the MDA measures. The results show that these measures are sometimes more closely linked to the target than the aM and rM, This is particularly the case when analysing the distributions of the RD map.

Looking at the measure-target scatter plots of this metric available in Appendix A, Figure A.3, we can clearly see the link between the *6-minute walk test* results and the measure's values. In particular, we note that an increase in the SCD, DTW, MCD or in the *Jensen-Shannon distance* seems to be linked to reduced brain health. This last result is logical as an increase in *Jensen-Shannon distance*, MCD or DTW means greater dissimilarities between the distributions of the metric within the contralesional and ipsilesional corticospinal tract. As an example, we can examine the Figure 4.1 which displays the smoothed distributions of the RD map of the same patient (20_07_01) a few days after the stroke and after its rehabilitation treatment. For the first plot, we see that the distribution of the RD over the damaged side of the CST has been slightly shifted to the left compared to the healthy side of the brain. This could be the effect of de-myelination of the axons in the region or the reduction of the thickness of axon bundles. After rehabilitation, the ipsilesional distribution re-aligns to the contralesional one, probably indicating the healing of the neurons as the results of the clinical test also improved. Here, this

re-alignment is reflected by the diminution of the *Jensen-Shannon distance* between the two curves.

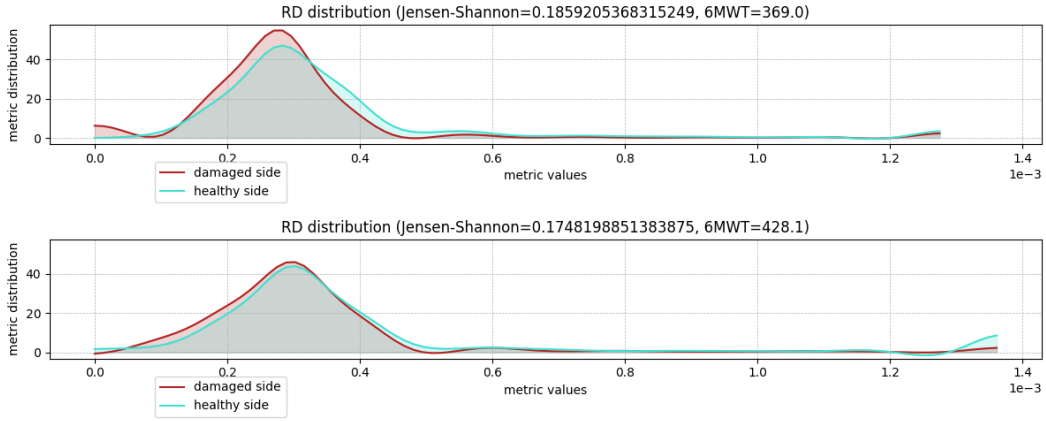

Figure 4.1: RD distributions of data 20_07_01_E0 and 20_07_01_E2

On the other hand, the *SCD* indicates if the differences between the two distributions are localized or spread over the distributions. For example, looking at the Figure 4.2 displaying the RD distributions of the patient 20_12_02 at the beginning and at the end of his rehabilitation, we observe that a gap between the ipsilesional and contralesional distributions is being filled throughout the rehabilitation, which reduces the value of the *SCD*. Also, we can see that the peak of the ipsilesional distribution is not entirely moving but is rather collapsing with a part of the density sliding to the right while the other remains in place. This could indicate that the volume of the distribution which does not move at all represents a part of the neurons that are definitely dead and not healing. If that is the case, this can be a good indication that the patient will not fully recover his or her abilities and that the treatment may focus on handling long-term disability.

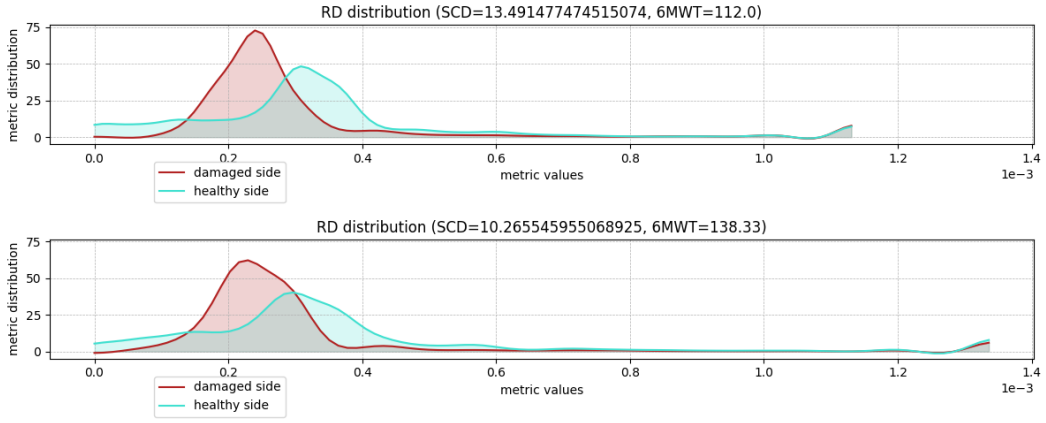

Figure 4.2: RD distributions of data 20_12_02_E0 and 20_12_02_E3

In terms of redundancy between the information provided by MDA and more traditional measures, we can clearly see that some of the information is overlapping across the different measures. That said, this redundancy is relative and not recurrent for all the measures and all

the metrics. A correlation coefficient of 0.6 certainly indicates a tendency between two features but it does not ensure that the two features cannot bring exclusive information about the target. The level of redundancy does not appear to be sufficient to completely abandon one feature in favour of another.

4.1.2 The MDA for the DIAMOND-enhanced DTI metrics

As expected, the results of the analysis of DIAMOND-enhanced DTI metrics are similar to those obtained for true DTI. It shows, once again, that some significant links appear to be possible between MDA measurements and patients' brain health. However, it is interesting to note that the calculation of the metrics by DIAMOND seems to improve the results of the aM and the rM, thus reducing the relative effectiveness of the other measures. The example that best illustrates this phenomenon is that of *radial diffusivity*, for which the aM, rM and CCD measurements become the most closely related to the target.

The analysis of the measure-target graphs (Appendix B, Figure B.4) for the *DIAMOND-enhanced mean diffusivity* provides interesting data to discuss the link between the results of the 6-minute walk test and the measures of aM, rM and CCD. We can see that an increase in aM or rM seems to be linked to better results in the clinical test. In fact, this result is surprising. Indeed, an increase in the *mean ratio* or *mean asymmetry* reflects a shift to the right of the mean of the metric over the ipsilesional corticospinal tract when compared to the contralesional one. Yet, a higher mean diffusivity is often induced by cellular death, which is not consistent with better clinical results. This is even more surprising given that analysis of the MD graph of the DTI in Figure A.4 does not indicate this type of relationship or that is at least much less pronounced. The same counter-intuitive relation is observed for the CCD, knowing that the CCD is inversely proportional to aM or rM. In contrast, results for the *DIAMOND-enhanced radial diffusivity* seem to be totally coherent.

For DIAMOND, MDA measures once again seem to contain at least some of the information already represented by aM and rM, particularly for the FA and RD metrics. In these two cases, the relevance of using MDA in conjunction with traditional measures could be questionable.

4.1.3 The MDA for the NODDI metrics

Overall, The results of the NODDI metrics analysis are much less fruitful, with little or no appreciable link between the metrics and the target, for the MDA measures as well as for the rM and the aM. The only metric of interest is the *orientation dispersion index*. For the latter, we find a strong relationship between the values of the aM, the rM, the CCD and, to a lesser extent, the *Jensen-Shannon distance* with those of the target. This finding leads us to believe that, for this metric, the traditional measures should prevail over those of the MDA. However, we note that the *Jensen-Shannon distance* seems to have very little correlation with rM and aM. This measure may therefore provide us with additional information related to the patient's brain health.

4.1.4 The MDA for the MF metrics

Because of the way it works, the MF diffusion model is the model that provides the best biological interpretation based on its metrics. It is therefore surprising to note that it is also the model with the least statistical significance in terms of establishing dependency links between the diffusion image and clinical test results of a patient. In fact, of the 3 metrics analysed for MF, only the *cerebrospinal fluid volume fraction* shows a statistically significant relationship. For this metric, there seems to be a link between the CCD, the DTW, the Jensen-Shannon distance and the target. Since traditional measures do not seem to benefit from this link, it is tempting to conclude that MDA measures find their full utility here. However, the mutual information also indicates that a non-linear relationship might exist between the aM, the rM and the target, forcing us to temper our analysis.

4.1.5 Summary

To sum up, the *metric distribution analysis* globally demonstrates that it can be used to reflect some of the microstructural transformations undergone by the brain after a stroke. Although MDA has proved its worth, it is not intended to completely replace more traditional measures such as *mean asymmetry* or the *mean ratio*. In fact, we can see that the effectiveness of these measures largely depends on the metrics analysed. While for some of them, MDA measures seem more appropriate, for others, traditional measures are more suitable. That is why it seems more advantageous to seek to obtain information on the patient's state of health through the simultaneous utilisation of these measures.

All the measures we presented previously are destined to compare both sides of the brain by evaluating some kind of distances between their respective metric distribution curves. Therefore, giving a biological interpretation of these measures only based on their value might be a rough exercise. However, visual analyses of the distribution curves provided by the MDA can be qualitatively useful and are more open to interpretation. The reading of these curves, as it has been done with Figures 4.1 and 4.2, could thus facilitate to highlight links between the distribution of a metric over a brain area and the microstructural reality of the patient.

4.2 Limitations

In addition to the limitations already discussed and specific to the diffusion models, our study also has some limitations due to the method and the data used. A review of these limitations is therefore necessary to contrast our conclusions.

Firstly, our study is limited to the analysis of diffusion metrics in the corticospinal tract only. Although this region of the brain is of key importance in voluntary motor control, it acts in conjunction with other areas that are not studied here. Limiting the study to the CST means missing out on potentially important changes in other areas of the brain that have an impact on patients' motor abilities, thus weakening our predictive possibilities. Moreover, looking at our data, we can see that, for many patients, the area affected by the stroke is not at all completely contained within the corticospinal tract.

Regarding our data, to increase the size of our sample, we used images of the same patients at different stages of rehabilitation as if they were different people. There are several ways in which this procedure could lead to bias in our results. On the one hand, it could induce a form of redundancy in the data if the acquisitions cannot really be considered as independent. On the other hand, the use of data collected at different stages of rehabilitation could introduce unwanted additional variance.

Furthermore, we use the results of the *6-minute walk test* as the only reference values for indicating patients' state of health. While this test has been shown to be effective in reflecting a patient's motor skills after a stroke, using this test alone does not allow us to accurately characterise a patient's overall brain health. As a result, the target feature used for our analysis is far from optimal, making our results less significant.

Another limitation of our study is the masking carried out by ElikoPy and the registration of the atlases. The masks generated by Elikopy, which were supposed to isolate the white matter from the rest of the tissues, were often too lax and did not completely remove the cranial cavity. As a result, our data was contaminated by a number of imperfections that somewhat affected our metric distributions.

We also noticed that the registration of the atlases and the projection of the regions of interest onto the patient space frequently led to a shrinking of the regions in question. This effect is undoubtedly due to the changes in the image caused by the stroke and affecting the registration. It is therefore highly likely that the selection of the corticospinal tract is affected and introduces a bias in the distribution of our metrics.

Finally, we limited ourselves to a univariate study of the relationships between the various features and the results of the *6-minute walk test*. We therefore have no idea of the implications that might result from linking several features in the search for relationships with the target. Furthermore, the only real statistically significant evidence of relationships between the target and certain features is based on a test designed to detect a form of linear dependency. It should also be remembered that although mutual information is an effective indicator, it does not guarantee by itself the existence of dependency relationships. As a result, some of the relationships we highlight in our conclusions are not fundamentally based on tangible statistics.

4.3 Future perspectives

As the 4.1.5 section suggests, there is still considerable room for improvement in our study. We can suggest a few areas for improvement.

In particular, we could extend the study to other brain regions involved in motor control. For example, the cerebellum, the superior longitudinal fasciculus or the corpus callosum could be included in the study. We know that these areas of the brain are involved in the voluntary control of movement, and their study in the context of stroke has already produced convincing results [51]. We could also further subdivide these areas in order to refine even more our study.

By continuing to talk about the regions of interest, we could also overcome the problems of registration and masking during the preprocessing phase. The cost in time would certainly be significant, but ideally, patient-by-patient inspection and hand-drawing of regions of interest and masks would limit the negative effects of automated delineation and improve by the same the significance of our results.

In order to eliminate the problems mentioned in relation to the use of images of the same patient at different stages of rehabilitation, we could completely revise the method of our analysis. For example, it would be possible to study the clinical and microstructural evolution of patients on an individual basis, rather than studying the state of their brain microstructure at specific times. We would then seek to link changes in the results of the *6-minute walk test* with changes in MDA during rehabilitation. For this study to have sufficient statistical weight, however, it would be preferable to expand our database and gather information on when images were acquired during the rehabilitation. In order to improve the comparability of our results, we could also imagine developing a reference pipeline for the study of MDA.

Furthermore, there is undoubtedly a lot of work to be done on the qualitative analysis of distribution curves. While for the time being we have noticed a few aberrations in our analyses, probably due to the limitations mentioned above, we have also seen that it is possible to draw quite relevant interpretations from these graphs.

A final example of an area to explore would be the multivariate study of the MDA. Even though it may be complicated to statistically prove a multivariate dependency between some features and a target, surrogate models could provide interesting results. Not satisfied with finding individual relationships between clinical tests and microstructural metrics, a multivariate study would be a further step towards establishing complex predictive models determining a patient's condition on the basis of diffusion images. We could then consider models based on the principle of the Multi Layer Perceptron (MLP), the K-Nearest Neighbours (KNN) or others that exploit the full potential of the information contained in the MDA of a diffusion image.

Conclusion

The characterisation of brain microstructures using diffusion imaging techniques is a prolific field of research. The associated results enable complex analysis of cerebral pathologies such as stroke, making a major contribution to the development of a better understanding of their mechanisms and improved patient care. This Master's thesis follows in the footsteps of these studies.

As we immersed ourselves in the literature on stroke and the use of diffusion imaging to characterise it, it was striking that we were still only scratching the surface of what was possible. That is how the new method of the MDA came into being. This method, based on distributions of metrics within targeted areas of interest, aims to exploit even further the potential of diffusion models and to use it to better understand stroke pathology.

To assess its validity, we described it in great detail. And so was explained step by step our method; Describing how the dMRI images were pre-processed and the diffusion metrics computed using the ElikoPy pipeline; How we had to register atlases onto the patient's space to identify the corticospinal tract; And explaining the analyses conducted on the metrics and the statistical validation of the results.

Very different results were obtained in the characterisation of our new method, depending on the metrics analysed. While results for the MF metrics did not show sufficient statistical significance to assess the use of the MDA, DTI, DIAMOND and, to a lesser extent, NODDI metrics revealed its relevance in stroke characterization. We saw that significant relationships could be deducted between some of the proposed measures and the patients' psychomotor skills. On top of its quantitative applications, the MDA also revealed its potential in a more qualitative way through the visual analysis of the metric distribution curves, enabling the study of patients' brain microstructures.

It is true that our discussions also revealed certain limitations to our study. Some are due to certain choices in the processing of our data and others are inherently linked to their very nature. Yet, each of these limitations provided an opportunity to explore new avenues of thought and to envisage certain improvements to be made in the future.

Finally, based on the results of this study and on our discussions, we can conclude that the joint use of traditional methods and MDA would certainly be beneficial to a better understanding of cerebral pathologies including stroke and that this thesis further opens up the field of possibilities in the analysis of diffusion images, paving the way for promising future research.

References

[1] Woodruff et al. *Pathophysiology, treatment, and animal and cellular models of human ischemic stroke*. Molecular Neurodegeneration 2011 6:11.

[2] Qin, C., Yang, S., Chu, YH. et al. *Signaling pathways involved in ischemic stroke: molecular mechanisms and therapeutic interventions*. Sig Transduct Target Ther 7, 215 (2022). https://doi.org/10.1038/s41392-022-01064-1

[3] Donkor ES. *Stroke in the 21st Century: A Snapshot of the Burden, Epidemiology, and Quality of Life*. Stroke Res Treat. 2018;2018:3238165. Published 2018 Nov 27. doi:10.1155/2018/3238165

[4] M Lipton. *Introducing MRI: Diffusion Imaging (49 of 56)*. Albert Einstein College of Medicine. Published on YouTube in September 2014 and consulted in September 2022.

[5] Doctor Klioze.*Diffusion Weighted Imaging*. On YouTube, published in December 2014, consulted in September 2022

[6] Cognitive Neuroscience Compendium.*MRI-based diffusion tensor imaging (DTI) and tractography*, on YouTube, published in April 2019, consulted in September 2022

[7] Koh SH, Park HH. *Neurogenesis in Stroke Recovery*. Transl Stroke Res. 2017;8(1):3-13. doi:10.1007/s12975-016-0460-z

[8] Anonymous. *STROKE Causes and Risk Factors*. National Heart, Lung and Blood Institute. On https://www.nhlbi.nih.gov/health/stroke/causes. Last updated on May 26, 2023. Consulted in June 2023.

[9] Ralph L. Sacco et al. *Risk Factors*. Stroke Vol. 28 No. 7. https://doi.org/10.1161/01.STR.28.7.1507Stroke.1997;28:1507âĂŞ1517

[10] Jin R, Yang G, Li G. *Inflammatory mechanisms in ischemic stroke: role of inflammatory cells*. J Leukoc Biol. 2010;87(5):779-789. doi:10.1189/jlb.1109766

[11] Anonymous. *Stroke - Treatment*. United Kingdom National Health Service. On https://www.nhs.uk/conditions/stroke/treatment/. Last updated on September 13 2022. Consulted in October 2022.

[12] Bindawas, Saad M, and Vishal S Vennu. *Stroke rehabilitation. A call to action in Saudi Arabia*. Neurosciences (Riyadh, Saudi Arabia) vol. 21,4 (2016): 297-305. doi:10.17712/nsj.2016.4.20160075

[13] Ryan J. Felling, Hongjun Song. *Epigenetic mechanisms of neuroplasticity and the implications for stroke recovery*. Experimental Neurology, Volume 268, Pages 37-45 (2015). ISSN 0014-4886. https://doi.org/10.1016/j.expneurol.2014.09.017

[14] E. V. Ekusheva, I. V. Damulin. *Post-Stroke Rehabilitation: Importance of Neuroplasticity and Sensorimotor Integration Processes*. Neuroscience and Behavioral Physiology, Vol. 45, No. 5, June, 2015. DOI 10.1007/s11055-015-0117-5.

[15] McDonnell, Michelle. *Action research arm test.* Aust J Physiother 54.3 (2008): 220.

[16] Sabrina Figueiredo, BSc. *Action Research Arm Test (ARAT).* Stroke Engine. On https://strokengine.ca/en/assessments/action-research-arm-test-arat/. Consulted in October 2022

[17] Chino, Naoichi, Sonoda, Shigeru, Domen, Kazuhisa, Saitoh, Eiichi, Kimura, Akio. (1996). *Stroke impairment assessment set (SIAS).* 10.1007/978-4-431-68461-9_3.

[18] Pohl, Patricia S., et al. *Influence of stroke-related impairments on performance in 6-minute walk test.* Journal of Rehabilitation Research & Development 39.4 (2002).

[19] Alexander, DC, Dyrby, TB, Nilsson, M, Zhang, H. *Imaging brain microstructure with diffusion MRI: practicality and applications*, NMR in Biomedicine. 2019; 32:e3841. https://doi.org/10.1002/nbm.3841

[20] Rith, K., & Schäfer, A. (1999). *The Mystery of Nucleon Spin.* Scientific American, 281(1), 58–63. http://www.jstor.org/stable/26058326

[21] B Kastler et al. *Comprendre l'IRM.* fr. OCLC: 1026755244. Elsevier Masson, 2011.

[22] thePIRL. *How MRI Works* (playlist), on YouTube, published between August 2018 and August 2022, consulted in September 2022

[23] F Peeters. LGBIO2050: *Medical Imaging, Chapter 4: Magnetic Resonance Imaging.* en. Ecole Polytechnique de Louvain, 2022.

[24] David C Preston. *Magnetic Resonance Imaging (MRI) of the Brain and Spine: Basics.* On https://case.edu/med/neurology/NR/MRI\%20Basics.htm. Published on 07/04/16. Consulted in November 2022.

[25] M Lipton. *Introducing MRI: Artifacts (46 of 56).* Albert Einstein College of Medicine. Published on YouTube in September 2014 and consulted in September 2022.

[26] E. Milshteyn. *MRI Artifacts.* Why & How. 15 October 2020. Published on YouTube in October 2020 and consulted in September 2022.

[27] Stadler Alfred, Schima Wolfgang, Ba'ssalamah Ahmed, Kettenbach Joachim, Eisenhuber Edith. (2007). *Artifacts in body MR imaging: Their appearance and how to eliminate them.* European radiology. 17. 1242-55. 10.1007/s00330-006-0470-4.

[28] Anonymous. *MRI Artifacts.* Oncology Medical Physics. On https://oncologymedicalphysics.com/mri-artifacts/. Consulted in June 2023.

[29] F. Raskin.*Microstructural characterization of the corticospinal tract in stroke patients using advanced diffusion Magnetic Resonance Imaging techniques.* Ecole polytechnique de Louvain, Université catholique de Louvain, 2022. Prom. : Macq, Benoît ; Dricot, Laurence. http://hdl.handle.net/2078.1/thesis:37927

[30] Soares Jose, Marques Paulo, Alves Victor, Sousa Nuno. *A hitchhiker's guide to diffusion tensor imaging.* Frontiers in Neuroscience, Volume 7 (2013). https://www.frontiersin.org/articles/10.3389/fnins.2013.00031. 10.3389/fnins.2013.00031

[31] Doctor Klioze. *Diffusion Tensor Imaging (DTI).* On YouTube, published in February 2019, consulted in September 2022

[32] Grover, Vijay P B et al. *Magnetic Resonance Imaging: Principles and Techniques: Lessons for Clinicians.* Journal of clinical and experimental hepatology vol. 5,3 (2015): 246-55. doi:10.1016/j.jceh.2015.08.001

[33] Huisman, T A G M. *Diffusion-weighted and diffusion tensor imaging of the brain, made easy.* Cancer imaging : the official publication of the International Cancer Imaging Society vol. 10 Spec no A,1A S163-71. 4 Oct. 2010, doi:10.1102/1470-7330.2010.9023

[34] Do Tromp. *DTI Scalars (FA, MD, AD, RD) - How do they relate to brain structure?.* On `https://www.diffusion-imaging.com/2013/01/relation-between-neural-microstructure.html`. Consulted in October 2022.

[35] Rutger Fick, Marco Pizzolato, Demian Wassermann, Rachid Deriche. *Diffusion MRI Anisotropy: Modeling, Analysis and Interpretation.* 2017. <hal-01468675>

[36] Hui Zhang et al. *NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain.* en. In: NeuroImage 61.4 (July 2012), pp. 1000–1016. issn: 10538119. doi: 10.1016/j.neuroimage.2012.03.072.

[37] Benoit Scherrer et al. *Characterizing brain tissue by assessment of the distribution of anisotropic microstructural environments in diffusion-compartment imaging (DIAMOND): Characterizing Brain Tissue with DIAMOND.* en. In: Magnetic Resonance in Medicine 76.3 (Sept. 2016), pp. 963–977. issn: 07403194. doi: 10.1002/mrm.25912.

[38] Gaëtan Rensonnet et al. *Generation of a dictionary of DW-MRI signals from arbitrary brain microstructure configurations.* Ecole polytechnique de Louvain, Université catholique de Louvain, 2015. Prom. : Macq, Benoît ; Taquet, Maxime.

[39] Gaëtan Rensonnet et al. *Towards microstructure fingerprinting: Estimation of tissue properties from a dictionary of Monte Carlo diffusion MRI simulations.* en. In: NeuroImage 184 (Jan. 2019), pp. 964–980. issn: 1053-8119. doi: 10 . 1016/j.neuroimage.2018.09.076.

[40] Quentin Dessain and Mathieu Simon. *Development of an image processing pipeline for large-scale in vivo population studies of the brain microstructure.* fr. Prom. : Macq, Benoît. Université catholique de Louvain, 2021.

[41] Quentin Dessain and Mathieu Simon. *ElikoPy - Documentation.* `https://elikopy.readthedocs.io/en/latest/index.html`. Consulted between November 2022 and Augustus 2023.

[42] Natali AL, Reddy V, Bordoni B. *Neuroanatomy, Corticospinal Cord Tract.* In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. Available from: `https://www.ncbi.nlm.nih.gov/books/NBK535423/` [Updated 2022 Aug 22].

[43] Anonymous. *Templates and Atlases included with FSL.* On `https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases`.

[44] Moura Luciana M. et al. *Diffusion Tensor Imaging Biomarkers to Predict Motor Outcomes in Stroke: A Narrative Review*. Frontiers in Neurology Vol. 10. 2019. doi : 10.3389/fneur.2019.00445.

[45] Anonymous. *Jensen–Shannon divergence*. On Wikipedia. `https://en.wikipedia.org/wiki/JensenâĂŞShannon_divergence`. Consulted in March 2023.

[46] Anonymous. *scipy.spatial.distance.jensenshannon*. On `https://docs.scipy.org/doc/scipy/reference/generated/scipy.spatial.distance.jensenshannon.html`. Consulted in March 2023.

[47] Thales Sehn Körting. *How DTW (Dynamic Time Warping) algorithm works*. On YouTube. `https://www.youtube.com/watch?v=_K1OsqCicBY`. Posted in September 2017. Consulted in March 2023.

[48] Anonymous. *Dynamic Time Warping (DTW)*. On `https://dtaidistance.readthedocs.io/en/latest/usage/dtw.html`. Consulted in March 2023.

[49] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, G.Varoquaux. *API design for machine learning software: experiences from the scikit-learn project*. ECML PKDD Workshop: Languages for Data Mining and Machine Learning (2013), pages 108-122.

[50] Michel Verleysen. *ELEC2870 - Machine learning: regression and dimensionality reduction - Feature selection*. Université catholique de Louvain, 2022.

[51] Hannanu, F.F., Naegele, B., Hommel, M. et al. *White matter tract disruption is associated with ipsilateral hand impairment in subacute stroke: a diffusion MRI study*. Neuroradiology 64, 1605–1615 (2022). `https://doi.org/10.1007/s00234-022-02927-8`

A Measures - target relations (DTI)

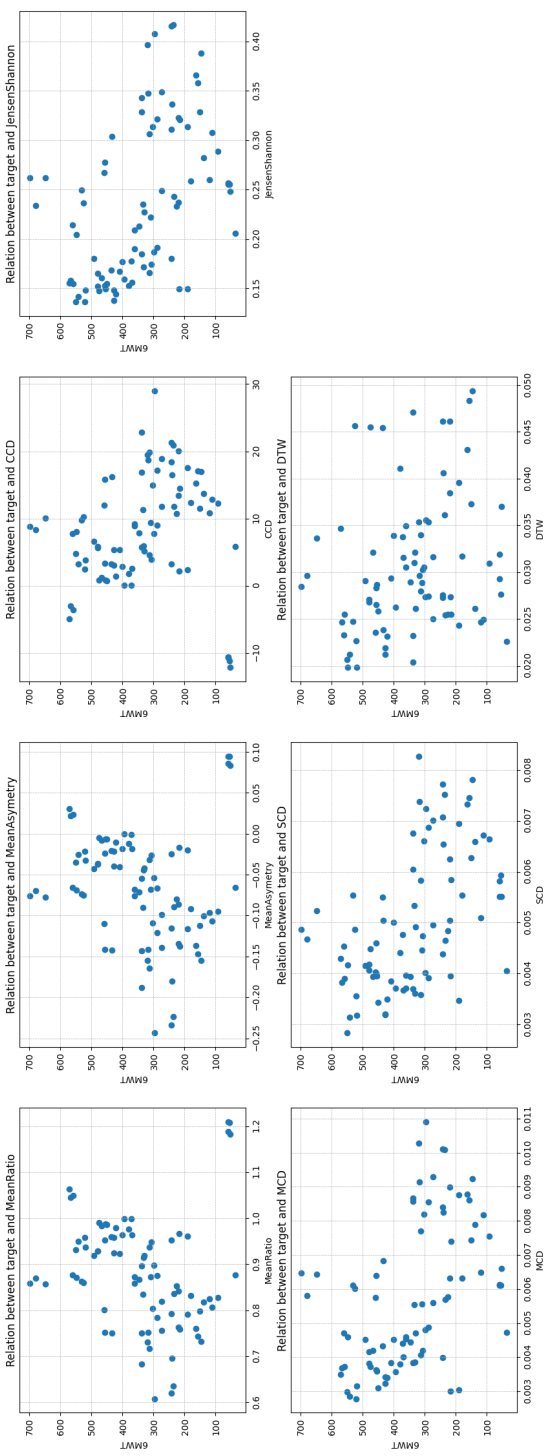

Figure A.1: Measure - 6MWT scatter plot for the FA metric.

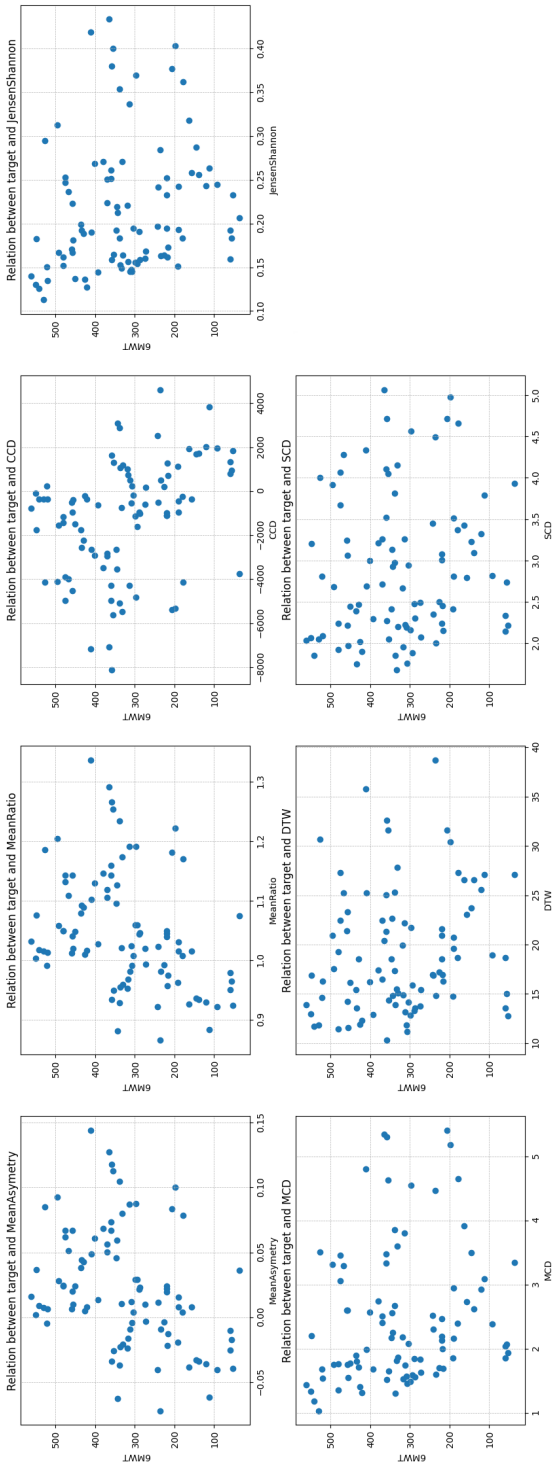

Figure A.2: Measure - 6MWT scatter plot for the AD metric.

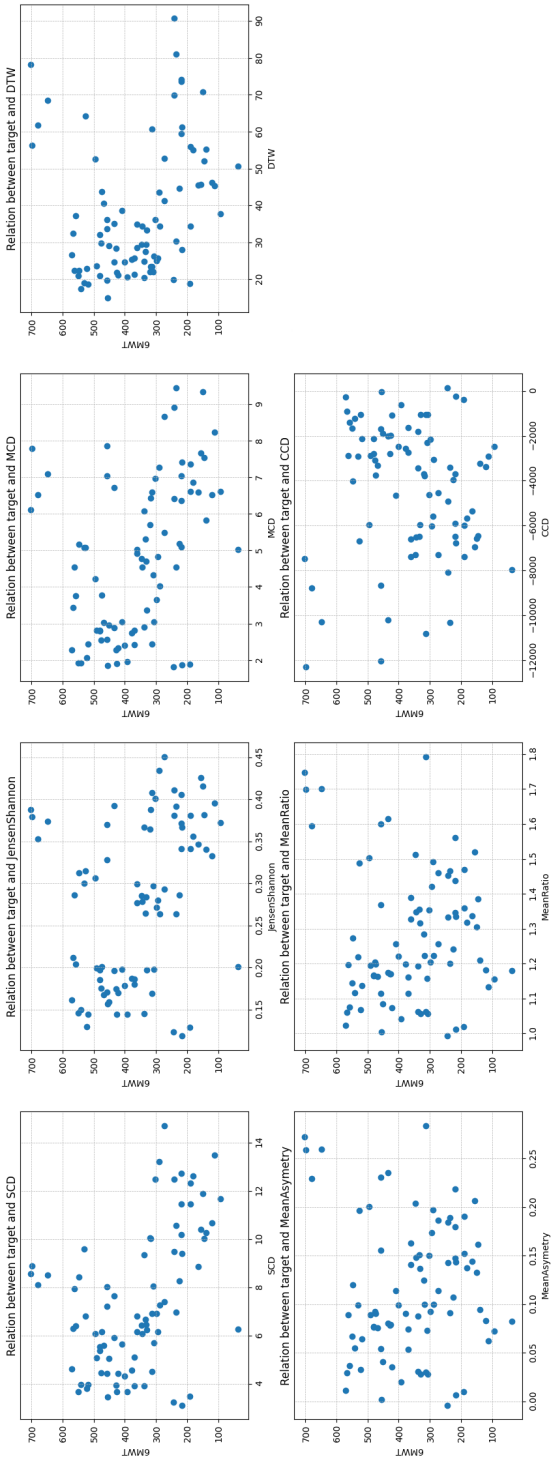

Figure A.3: Measure - 6MWT scatter plot for the RD metric.

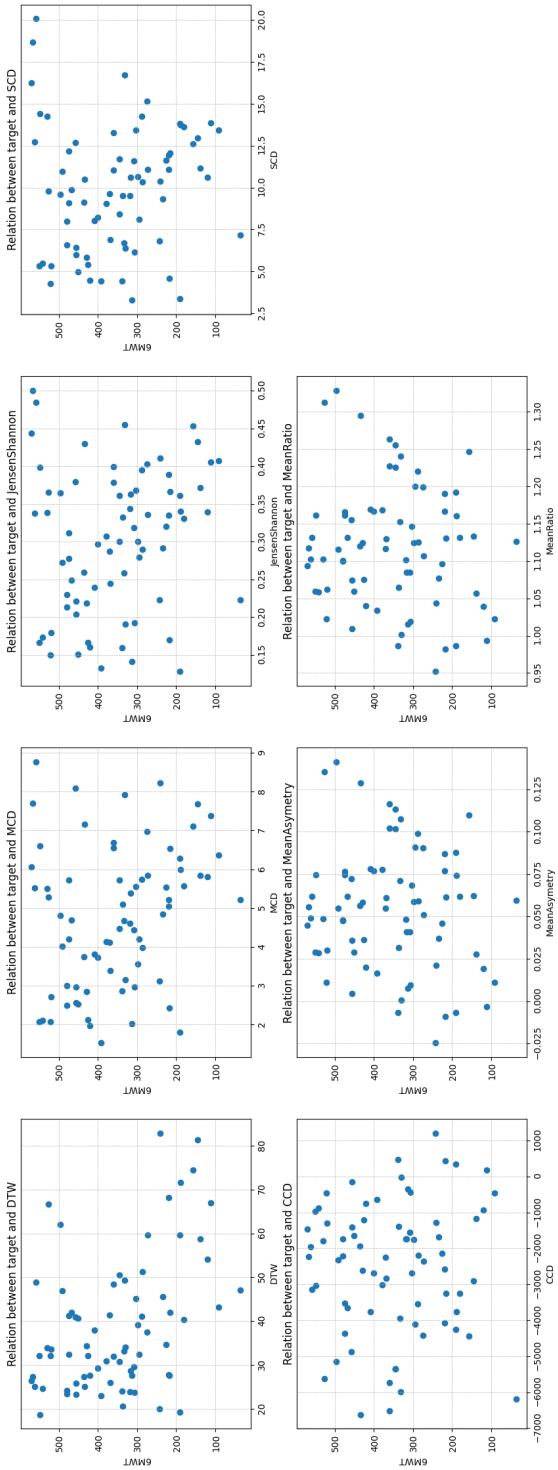

Figure A.4: Measure - 6MWT scatter plot for the MD metric.

B Measures - target relations (DIAMOND)

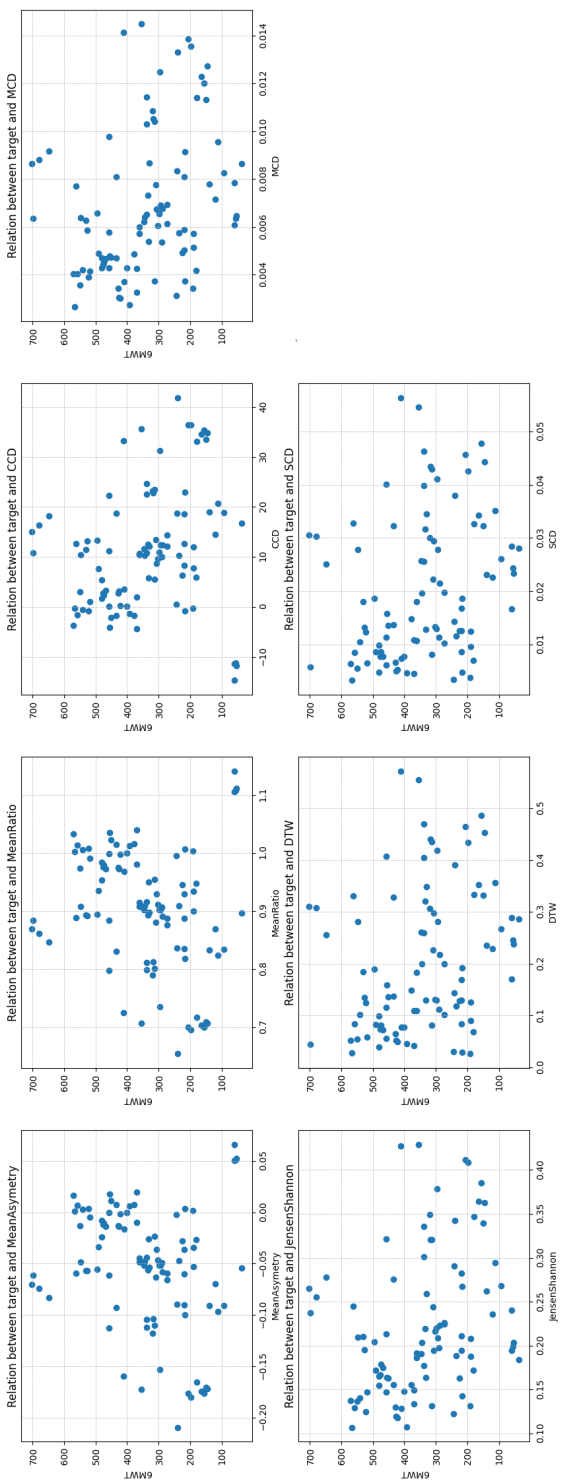

Figure B.1: Measure - 6MWT scatter plot for the DIAMOND-enhanced FA metric.

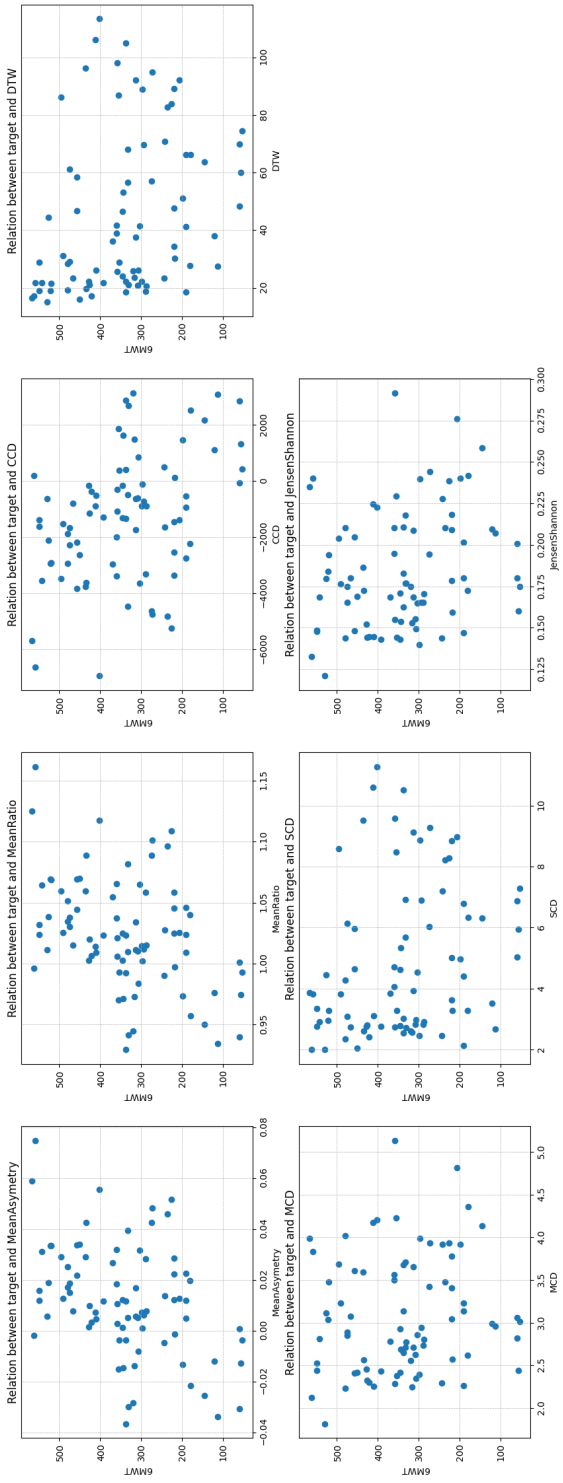

Figure B.2: Measure - 6MWT scatter plot for the DIAMOND-enhanced AD metric.

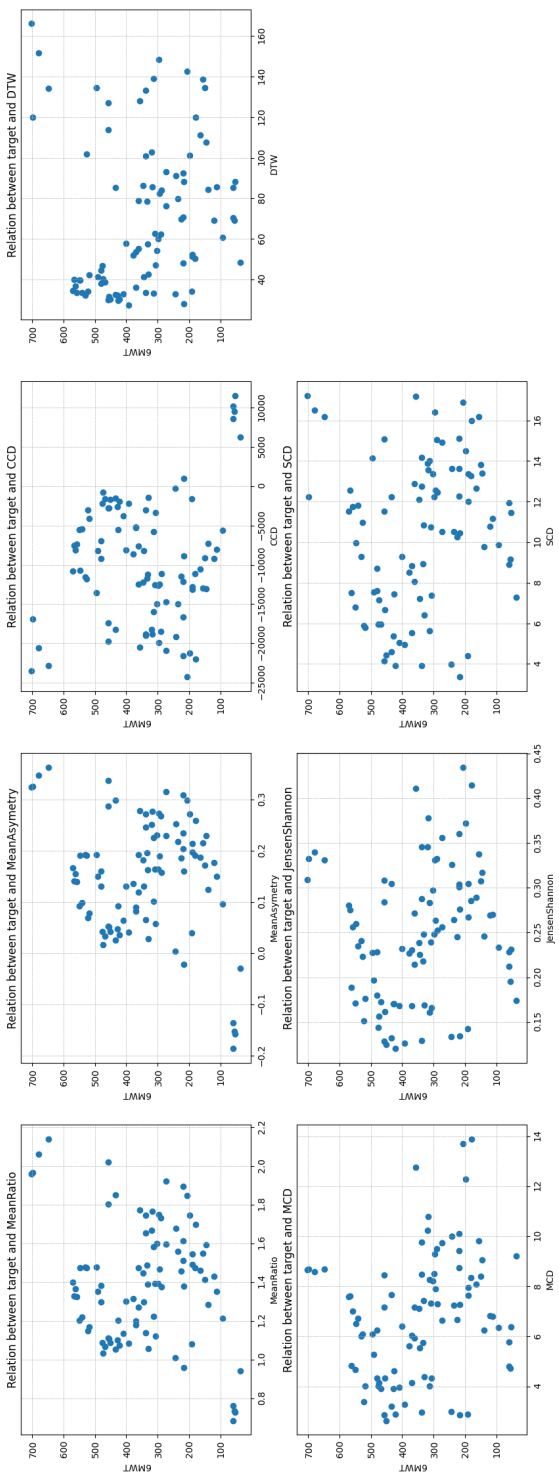

Figure B.3: Measure - 6MWT scatter plot for the DIAMOND-enhanced RD metric.

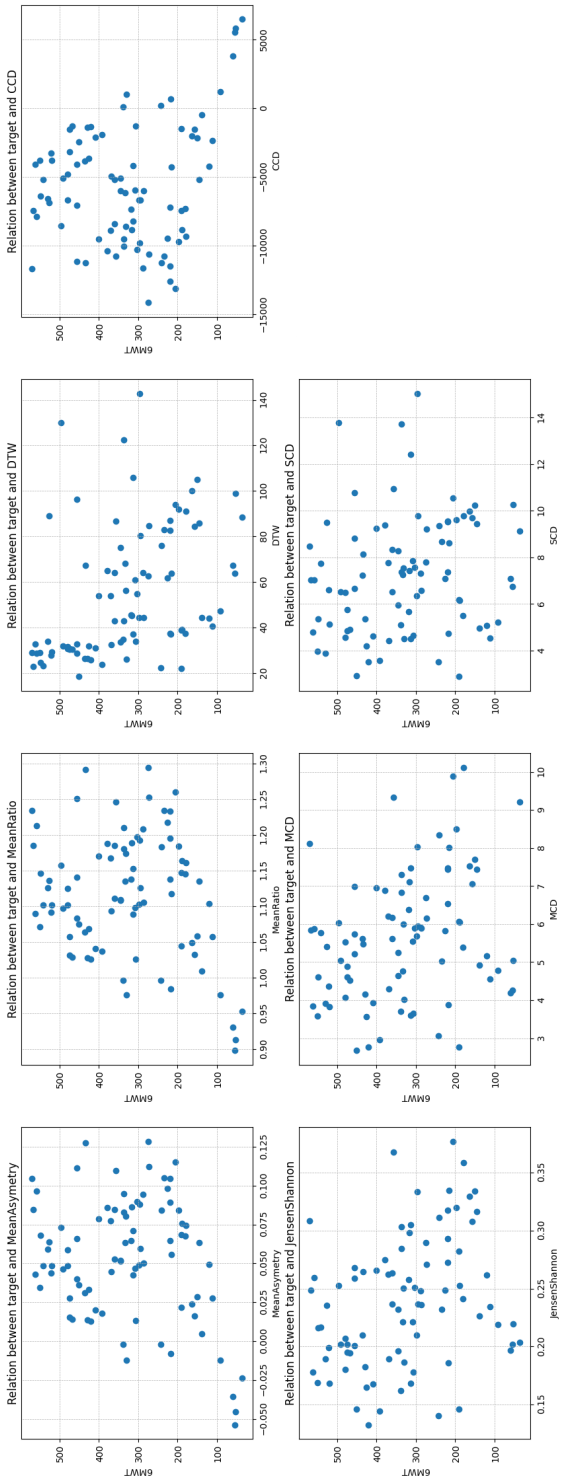

Figure B.4: Measure - 6MWT scatter plot for the DIAMOND-enhanced MD metric.

C Measures - target relations (MF)

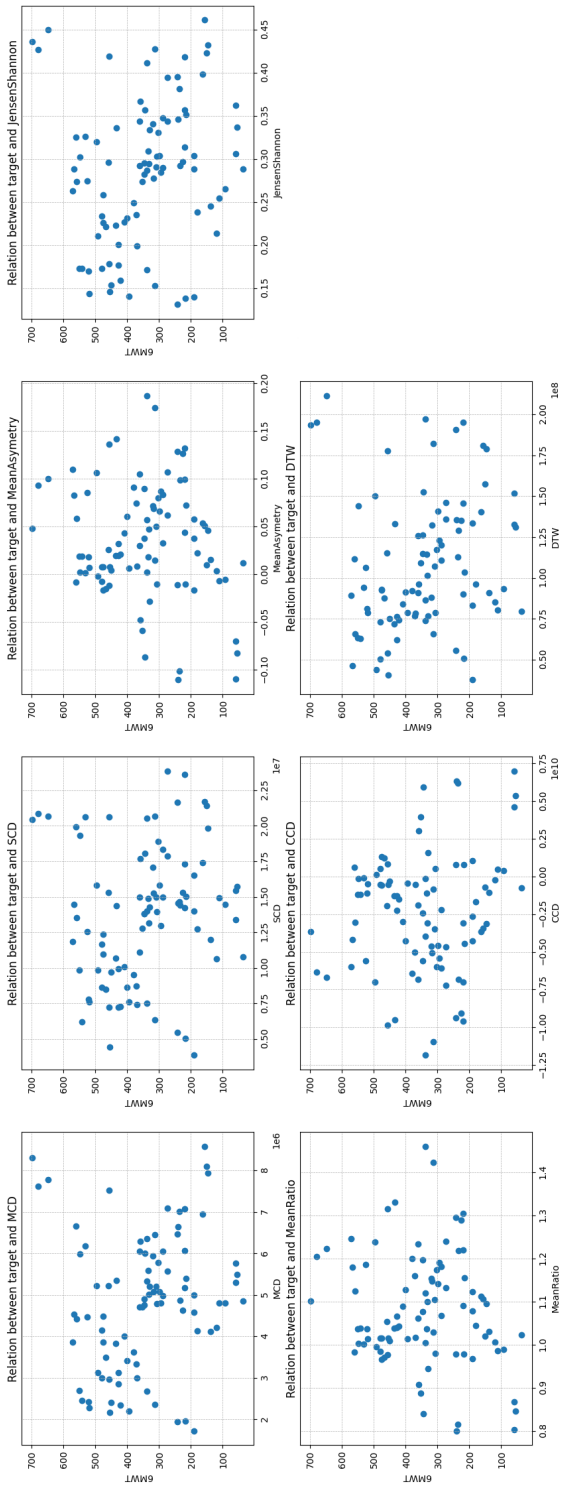

Figure C.1: Measure - 6MWT scatter plot for the DIFF_ex_tot metric.

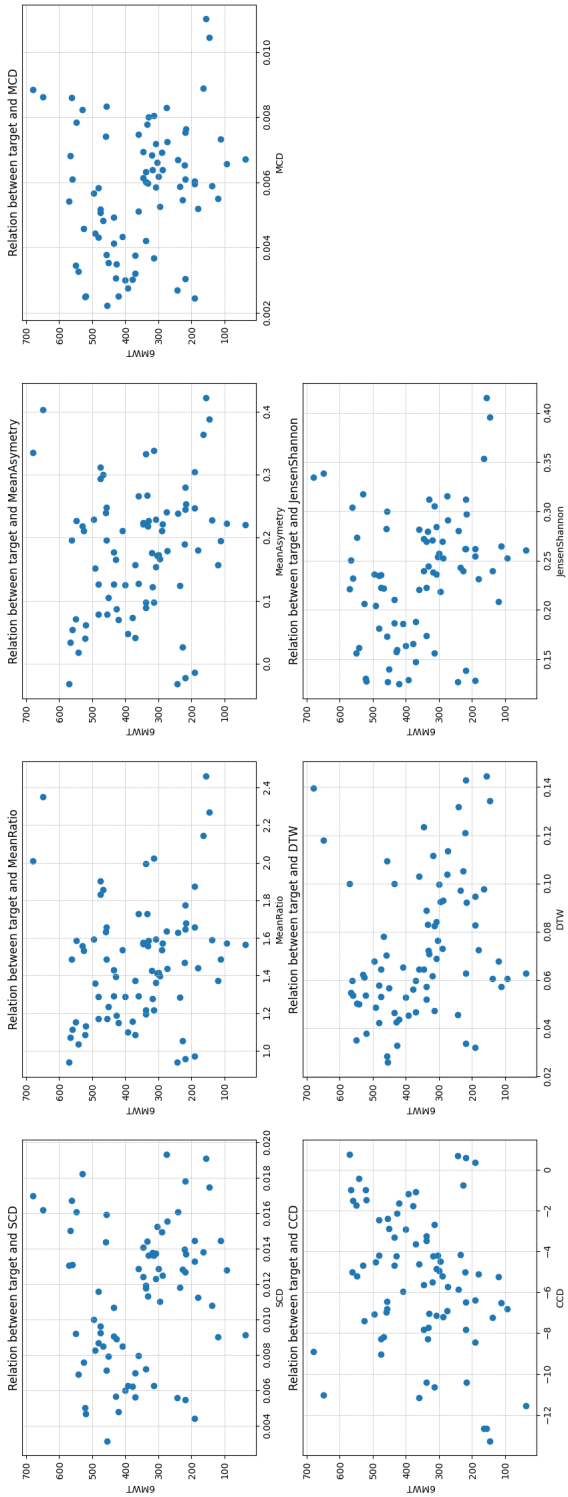

Figure C.2: Measure - 6MWT scatter plot for the frac_csf metric.

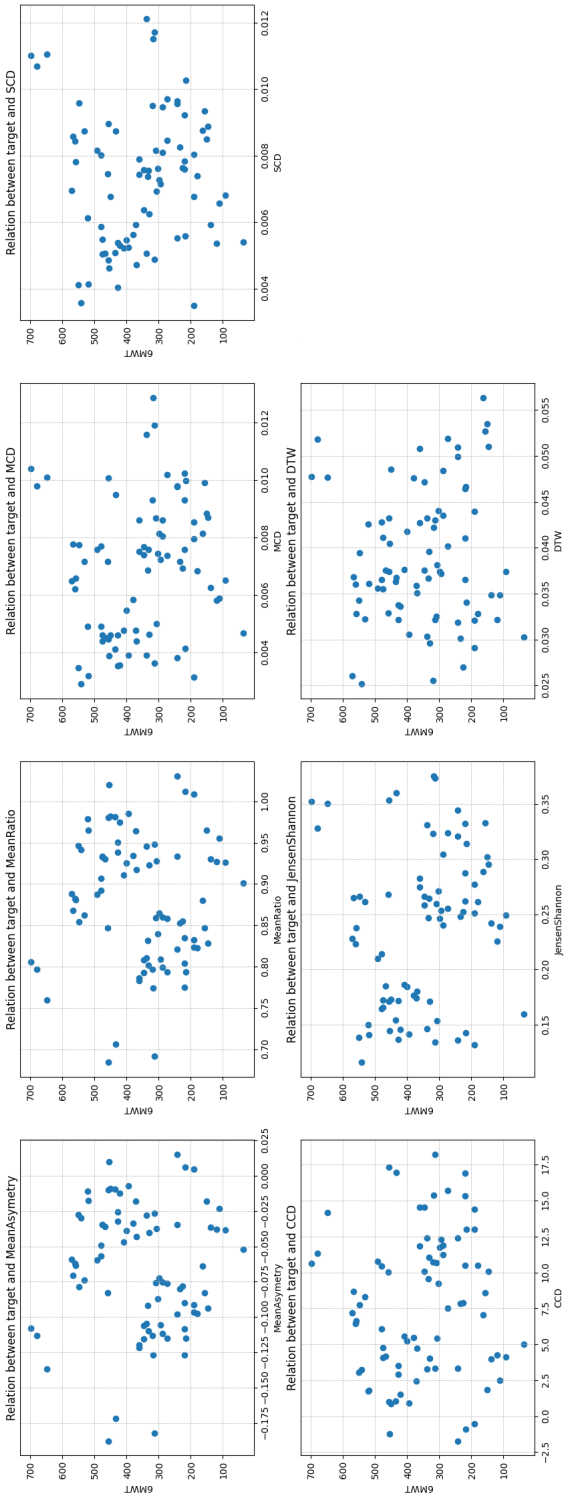

Figure C.3: Measure - 6MWT scatter plot for the fvf_tot metric.

D Measures - target relations (NODDI)

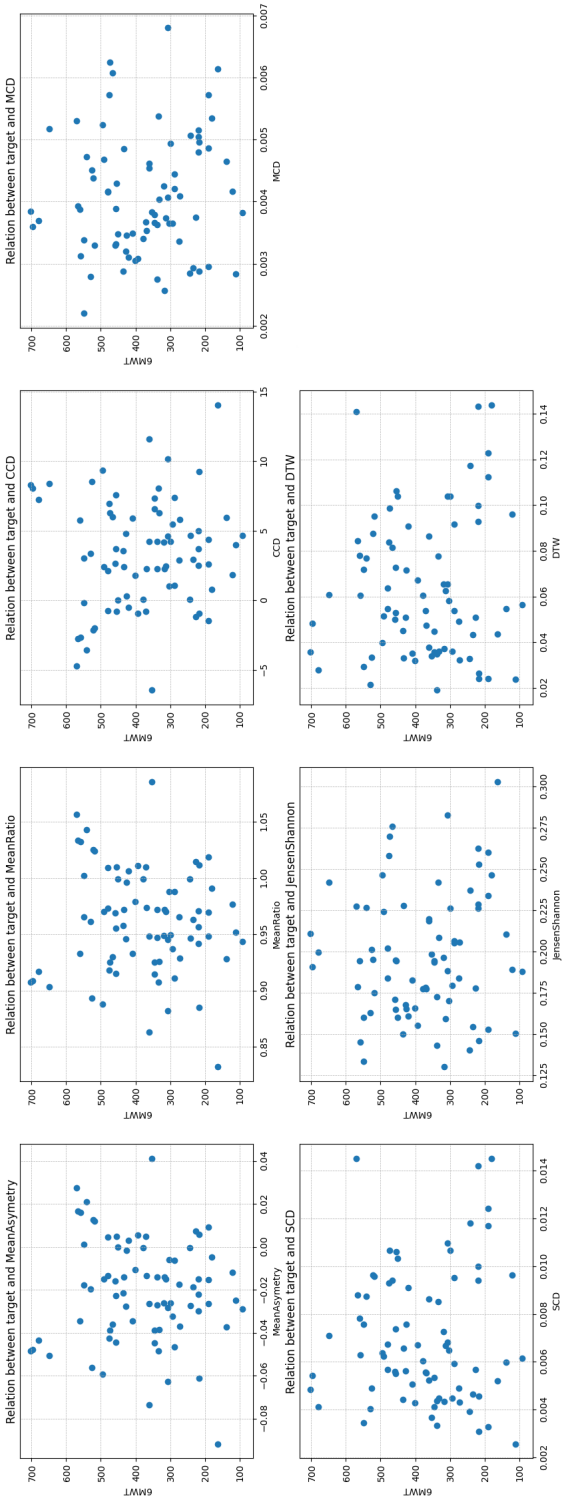

Figure D.1: Measure - 6MWT scatter plot for the fbundle metric.

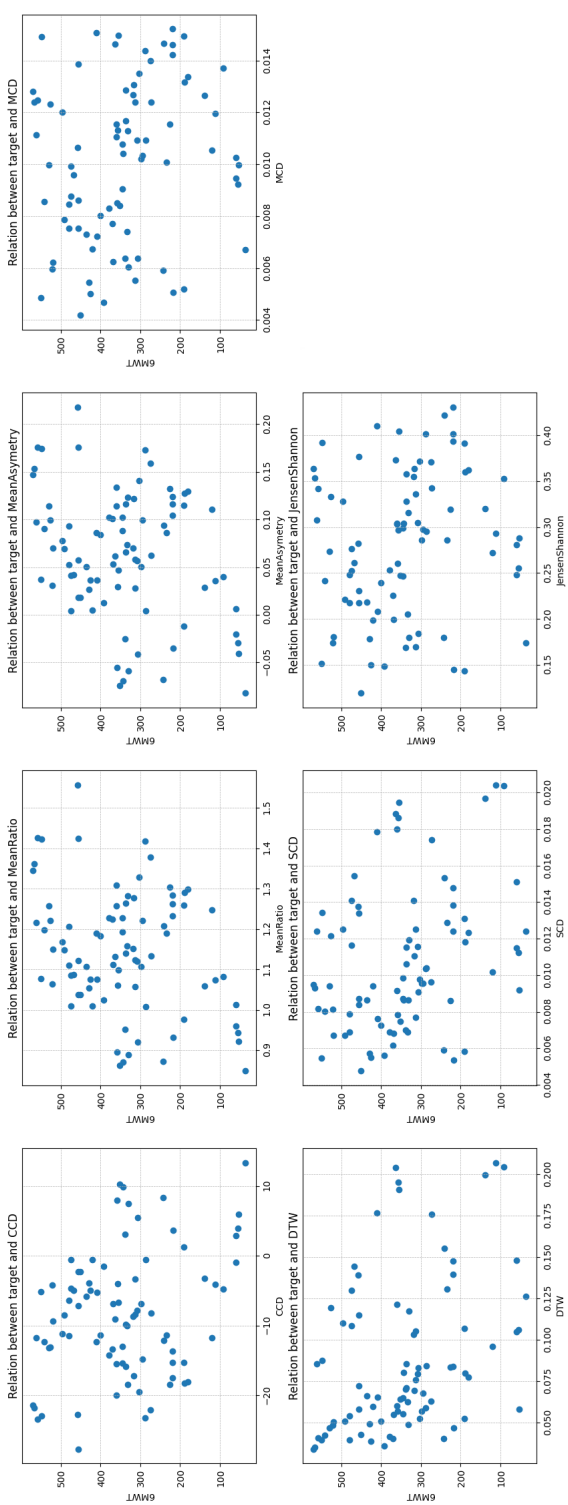

Figure D.2: Measure - 6MWT scatter plot for the fextra metric.

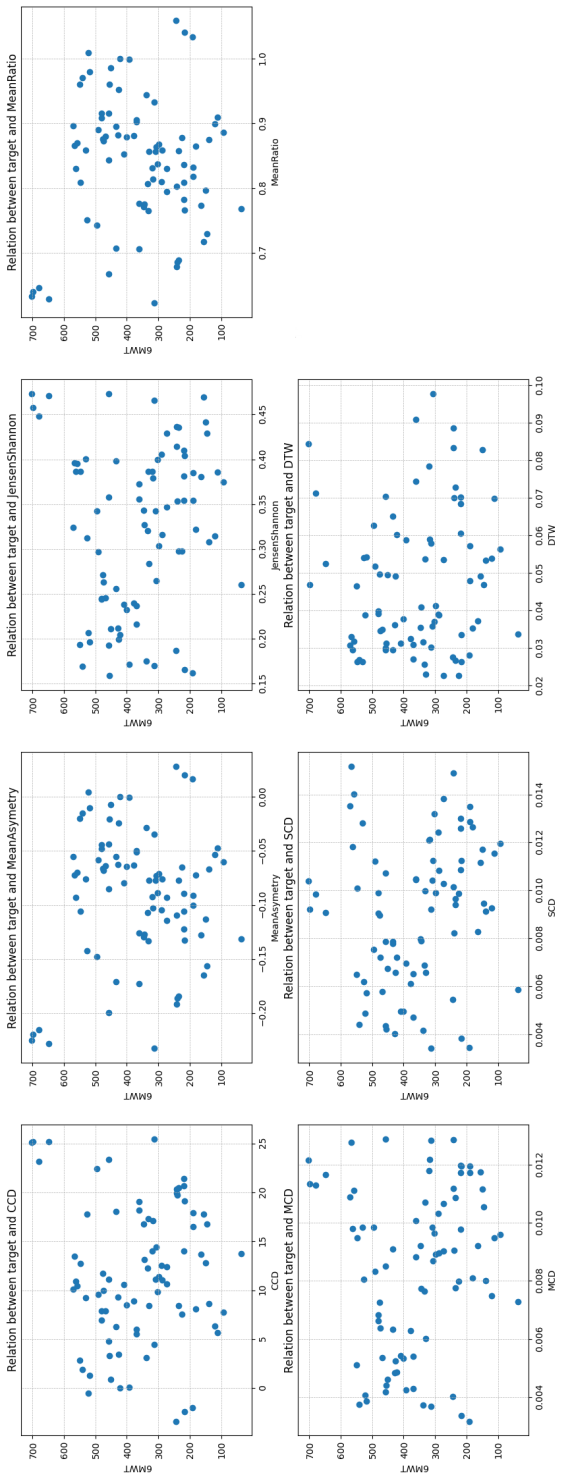

Figure D.3: Measure - 6MWT scatter plot for the fintra metric.

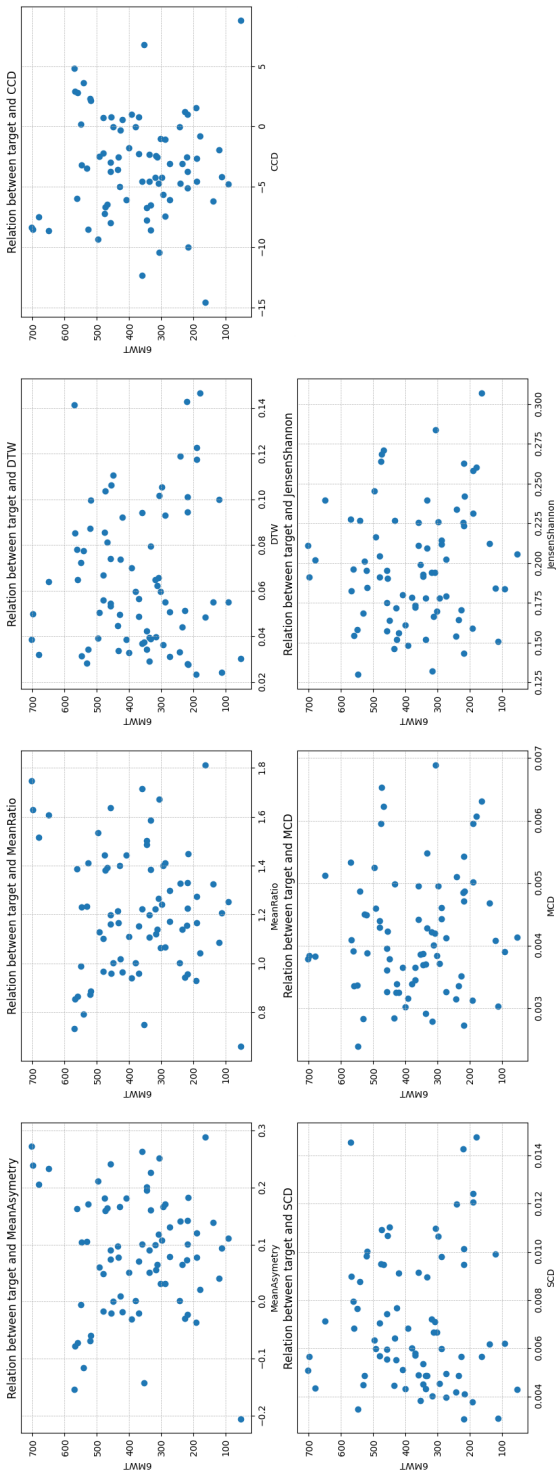

Figure D.4: Measure - 6MWT scatter plot for the fiso metric.

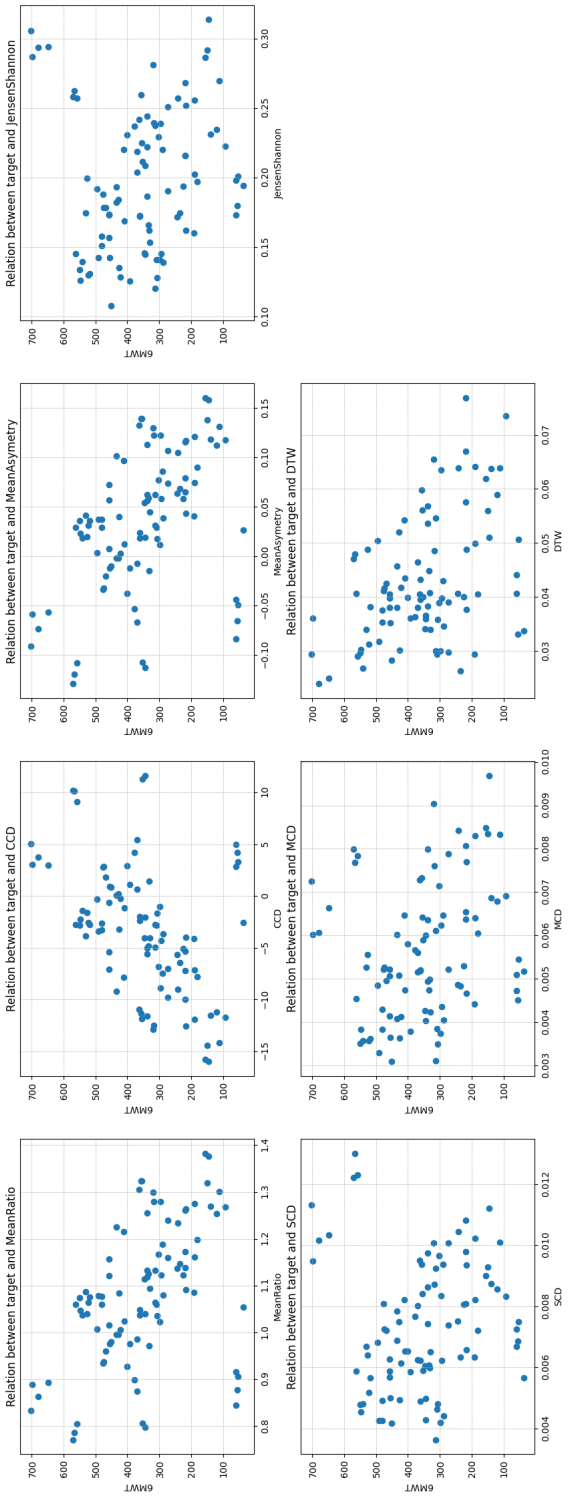

Figure D.5: Measure - 6MWT scatter plot for the ODI metric.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain
Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/epl**