UCLouvain

epl

École polytechnique de Louvain

Connectivity studies through MRI for alcohol dependent patients to identify biomarkers

Author: **Camille Dubois**
Supervisors: **Benoît Macq, Laurence Dricot**
Readers: **Nicolas Delinte, Mélissa Salavrakos**
Academic year 2022–2023
Master [120] in Mathematical Engineering

Abstract

Diffusion magnetic resonance imaging is a powerful tool used to study brain structures and connectivity. It has been widely employed in evaluating the impact of alcohol use disorder. This study focuses on investigating the relationship between structural connectivity changes observed during abstinence and the corresponding changes in behavioral measures. The research was conducted on a group of 73 patients to gain insights into biomarkers of behavioral recovery during abstinence by studying structural brain connectivity. The subjects underwent three scans: one on the first day of an abstinence program, a second 18 days later, and a third three months later. The connectome-based predictive model designed by Shen et al. was applied on structural brain connectivity data obtained from tractograms to quantify brain-behavior relationships. The behaviors studied are depression, obsessive-compulsive disorders, state and trait anxiety, and multidimensional fatigue, which are known to be affected by alcohol consumption. Obsessive-compulsive disorders measures were divided into cravings of alcohol, obsessions, and compulsions. A progressive evolution of the behavioral measures towards the controls has been observed during the first three months of abstinence. Moreover, the selected biomarkers revealed a predominance of the left hemisphere in predicting the evolution of behavioral measures during abstinence. Depression was primarily associated with parietal regions, while also involving the left lateral occipital white matter and the left transverse temporal white matter. Temporal and occipital regions played a significant role in state and trait anxiety. Biomarkers for multidimensional fatigue were predominantly located in the left temporal and parietal regions, with the left paracentral white matter also identified as influential. For cravings and compulsions, the left superior parietal cortex, brain stem, and left precuneus cortex were identified as biomarkers. Additionally, the right choroid plexus and right banks of the superior temporal sulcus emerged as biomarkers for cravings, whereas the right hippocampus and left transverse temporal white matter were associated with compulsions. In the case of obsessions, the brain stem, mid posterior corpus callosum, left superior temporal cortex, right insular cortex, and left thalamus proper were identified as biomarkers. These findings align with the known functions of the identified regions, as well as with the known regions altered by AUD or involved for the behaviors studied. Biomarkers of behavioral recovery in patients suffering from alcohol use disorder during abstinence can thus be identified.

Keywords: Alcoholism, Alcohol Use Disorder, Abstinence, MRI, Brain, Connectivity, Behavior, Biomarkers, CPM.

Acknowledgements

First and foremost, I express my gratefulness to my supervisors, Professor Benoît Macq and Professor Laurence Dricot, for the opportunity to work on such an exciting subject.

I would like to thank Mélissa Salavrakos and Philippe de Timary. Their expertise and valuable discussions have greatly contributed to my work.

I am deeply indebted to my assistants Nicolas Delinte and Manon Dausort, for their support, their answers to my numerous questions, their reactivity, valuable insights, comments and advice, and for the all the hours they gave me.

Lastly, I would like to express my profound gratitude to my family and friends who supported and encouraged me during my studies.

List of Figures

1.1 Spin and magnetic moment of a particle . . . 3
1.2 Effect of an external magnetic field B_0 on protons . . . 4
1.3 Effects of a RF impulsion on the magnetization vector $\vec{M}$ . . . 5
1.4 Effects of the different inhomogeneities on the time constant T_2 . . . 7
1.5 180° RF pulse to cancel the constant inhomogeneities linked to $\vec{B}_0$ . . . 8
1.6 Spin Echo sequence . . . 8
1.7 Influence of TR on the T_1-weighted contrast . . . 9
1.8 Influence of TE on the T_2-weighted contrast . . . 9
1.9 Proton density weighted contrast . . . 10
1.10 Pulsed Gradient Spin-Echo sequence and effects on the phase of the water molecules . . . 11
1.11 DTI Tensor depending on the type of diffusion . . . 13
1.12 Example of deterministic tractography and probabilistic tractography . 14

2.1 Examples of graphs with their adjacency matrix . . . 16

3.1 Worldwide alcohol consumption and alcohol use disorders . . . 17
3.2 Default Mode Network . . . 20
3.3 Cerebellum and frontal lobe . . . 22

4.1 Steps to reach the aim of this thesis . . . 29
4.2 Comparison of a scan before and after preprocessing with `ElikoPy`. . . 32
4.3 Wmparc map of a healthy participant. . . . 33
4.4 Example of registration. . . . 33
4.5 Illustration of the dilution phenomenon. . . . 34
4.6 Tractograms for an AUD participant and a control participant. . . . 35
4.7 A connectivity matrix before and after the sorting of the labels. . . . 36
4.8 A connectivity matrix without and with normalization. . . . 37
4.9 Schematic of connectome-based predictive modeling . . . 38
4.10 Evolution of the different behavior metrics for the AUD participants. . 39

5.1 Evolution of the different behavioral measures without the outliers and comparison with the controls' behavioral measures. . . . 48
5.2 Anterior, lateral, and superior views of the regions selected for BDI. . . 53
5.3 Anterior, lateral, and superior views of the regions selected for OCDS. . 54
5.4 Anterior, lateral, and superior views of the regions selected for Obsessions. 54
5.5 Anterior, lateral, and superior views of the regions selected for Compulsions. . . . 54

5.6 Anterior, lateral, and superior views of the regions selected for STAI. . 55
5.7 Anterior, lateral, and superior views of the regions selected for MFI. . . 55

6.1 Illustration of the different regions selected for each of the behaviors, and their connections to one another. 59

A.1 ElikoPy pipeline steps 84

C.1 Boxplots from which the outliers were identified. 86

List of Tables

3.1 Impact of AUD on the connectivity strength of different neural links . . 24

4.1 Correlations and p-values between the different behavior metrics, average absolute motion, and average relative motion with the participants. 40
4.2 Results of the application of CPM to predict motion 44

5.1 Selected sets of parameters with the smallest p-values for each behavior, and their prediction performance. 49
5.2 P-values obtained with the permutation testing for each behavior and model selected. 50
5.3 Number of significant edges identified for each of the models with the best prediction accuracies, and for each of the behaviors. 50
5.4 Number of edges linking the different hemispheres of the brain for each behavior. 51
5.5 Number of edges linking the different structures of the brain for each behavior. 51
5.6 Percentage of p-values below 0.05 and below the Bonferroni thresholds for the different behavioral measures and different models. 56

B.1 Number of labels of each parcellation map produced by the `recon-all` command. 85

D.1 Different sets of parameters used for CPM. 87
D.2 Continuation of table D.1: Different sets of parameters used for CPM. . 88

E.1 Correlations for the different sets of parameters for positive associations of the "Separate" model using partial correlation. 89
E.2 Correlations for the different sets of parameters for negative associations of the "Separate" model using partial correlation. 90
E.3 Correlations for the different sets of parameters for the "Joint" model using partial correlation. 91
E.4 Correlations for the different sets of parameters for the "Together" model using partial correlation. 92
E.5 p-values for the different sets of parameters for the positive associations of the "Separate" model using partial correlation. 93
E.6 p-values for the different sets of parameters for the negative associations of the "Separate" model using partial correlation. 94
E.7 p-values for the different sets of parameters for the "Joint" model using partial correlation. 95

E.8 p-values for the different sets of parameters for positive associations the "Together" model using partial correlation. 96

F.1 Selected connections for BDI, as well as the model in which they were selected. 97
F.2 Selected connections for OCDS, as well as the model in which they were selected. 100
F.3 Selected connections for Obsessions, as well as the model in which they were selected. 101
F.4 Selected connections for Obsessions, as well as the model in which they were selected. 103
F.5 Selected connections for STAI, as well as the model in which they were selected. 104
F.6 Selected connections for MFI, as well as the model in which they were selected. 107

Acronyms

ACC Anterior Cingulate Cortex.

AD Axial Diffusivity.

AIC Akaike Information Criterion.

AN Attention Network.

AUD Alcohol Use Disorder.

BDI Beck Depression Inventory.

BEARNI Brief Evaluation of Alcohol-Related Neuropsychological Impairment.

CPM Connectome-based Predictive Modeling.

CSF Cerebro Spinal Fluid.

DMN Default Mode Network.

DTI Diffusion Tension Imaging.

DWI Diffusion Weight Imaging.

ECN Executive Control Network.

FA Fractional Anisotropy.

FID Free Induction Decay.

GFE Frequency Encoding Gradient.

GPE Phase Encoding Gradient.

GSS Slice Selection Gradient.

ICQ Impaired Control Questionnaire.

MCC Middle Cingulate Cortex.

MD Mean Diffusivity.

MFI Multidimensional Fatigue Inventory.

MPPCA Marchenko-Pastur Principal Component Analysis.

MRI Magnetic Resonance Imaging.

MSE Mean Squared Error.

MSMT-CSD Multi-Shell Mutli-Tissue Constrained Spherical Deconvolution.

NAcc Nucleus Accumbens.

NMR Nuclear Magnetic Resonance.

OCDS Obsessive Compulsive Drinking Scale.

ODF Orientation Distribution Function.

PCA Principal Component Analysis.

PCC Posterior Cingulate Cortex.

RD Radial Diffusivity.

RF Radio Frequency.

RN Reward Network.

SADQ Severity of Alcohol Dependence Questionnaire.

SE Spin Echo.

SFG Superior frontal gyrus.

SIFT Spherical-deconvolution Informed Filtering of Tractograms.

SN Salience Network.

SsN Somatosensory Network.

STAI State-Trait Anxiety Trait.

TE Echo Time.

TR Repetition Time.

Contents

Abstract i
Acknowledgements ii
List of figures iii
List of tables iv
Acronyms vi

Introduction **1**

I State of the art **2**

1 Magnetic Resonance Imaging **3**
1.1 Imaging through Nuclear Magnetic Resonance 3
1.1.1 Magnetic resonance and Radio Frequency pulses 4
1.1.2 Relaxation 6
1.1.3 Spin Echo sequence 7
1.1.4 Image contrast 8
1.1.5 Acquisition 10
1.2 Diffusion Weight Imaging 11
1.2.1 Diffusion Tension Imaging (DTI) 12
1.2.2 Tractography 13

2 Graph Theory and Connectivity **15**
2.1 Graph Theory 15
2.2 Connectivity 15

3 Alcohol Use Disorder **17**
3.1 Definition 17
3.2 Impact of AUD 18
3.2.1 Biological impact 18
3.2.2 Behavioral and psychological impact 19
3.3 Abstinence and relapse 19
3.4 Analysis of connectivity and AUD 19
3.4.1 Default Mode Network (DMN) 20
3.4.2 Executive Control Network (ECN) 20
3.4.3 Salience Network (SN) 21
3.4.4 Attention network (AN) 21
3.4.5 Somatosensory network (SsN) 22
3.4.6 Reward network (RN) 22

3.4.7 Cerebellum . . . 22
3.4.8 Other neural links affected by AUD . . . 23
3.4.9 Summary . . . 24

II Analysis **28**

4 Methods **29**
4.1 Data collection . . . 30
4.1.1 Structural data . . . 30
4.1.2 Behavioral data . . . 30
4.2 Connectivity matrices . . . 31
4.2.1 Preprocessing . . . 31
4.2.2 Labeling . . . 32
4.2.3 Registration . . . 33
4.2.4 Structural connectivity . . . 34
4.3 Connectome-based Predictive Model . . . 37
4.3.1 Preliminary steps . . . 38
4.3.2 Model and parameters used . . . 40
4.3.3 Motion prediction . . . 44
4.3.4 Correcting for false positives . . . 45
4.4 Final models selection . . . 46
4.5 Extracting the biomarkers . . . 46

5 Results **47**
5.1 Behavioral changes during abstinence . . . 47
5.2 Best fitting CPM models . . . 48
5.2.1 Selected parameter sets . . . 48
5.2.2 Permutation testing . . . 49
5.3 Biomarkers . . . 50
5.3.1 Selected connections . . . 51
5.3.2 Regions as biomarkers . . . 53
5.4 Factors affecting CPM predictions . . . 55
5.4.1 Effects of the different parameters . . . 55

6 Discussion **58**
6.1 Biomarkers . . . 58
6.1.1 Connections . . . 58
6.1.2 Regions . . . 60
6.1.3 Comparison of OCDS, Obsessions, and Compulsions . . . 65
6.2 CPM . . . 66
6.3 Limitations . . . 67
6.4 Further perspectives . . . 68

Conclusion **69**

Bibliography **71**

Appendices

A ElikoPy **83**

B Number of labels by map produced by recon-all **85**

C Boxplots used to detect the outliers for the different behavior metrics **86**

D Different sets of parameters used for CPM **87**

E Results using partial correlation **89**
E.1 Correlations . . . 89
E.1.1 "Separate" Model – Positive edges . . . 89
E.1.2 "Separate" Model – Negative edges . . . 90
E.1.3 "Joint" Model . . . 91
E.1.4 "Together" Model . . . 92
E.2 *p*-values . . . 93
E.2.1 "Separate" Model – Positive edges . . . 93
E.2.2 "Separate" Model – Negative edges . . . 94
E.2.3 "Joint" Model . . . 95
E.2.4 "Together" Model . . . 96

F Connections selected with a threshold of 20 edges **97**
F.1 BDI . . . 97
F.2 OCDS . . . 100
F.3 Obsessions . . . 101
F.4 Compulsions . . . 103
F.5 STAI . . . 104
F.6 MFI . . . 107

G Assessing the impact of correcting for motion on the results **109**

Introduction

The goal of this thesis is to identify biomarkers of behavioral recovery in patients suffering from alcohol use disorder during abstinence by studying their brain connectivity.

Alcohol Use Disorder (AUD) is "a medical condition characterized by an impaired ability to stop or control alcohol use despite adverse social, occupational, or health consequences" [1]. It is responsible for 3 million deaths per year, which accounts for 5.3% of the global number of deaths per year [2]. It is thus a widely spread disease, affecting multiple age groups around the world.

Alcohol consumption not only affects mental health and behavior but also impacts brain microstructures. Those changes are linked, since the affected behaviors are governed by the brain networks impaired by alcohol consumption. Moreover, a period of abstinence reverses some of the damages caused by alcohol consumption on the brain, and influences the behaviors impacted by AUD.

Magnetic Resonance Imaging and diffusion Magnetic Resonance Imaging allow to study the brain structures and connectivity, which makes them great tools to quantify how they are damaged by alcohol. While there is abundant literature on alcohol and brain damage, less is known about the links between the structural connectivity changes observed during abstinence, and the evolution of behavioral measures during the same period.

We have taken a global approach, using the whole brain and not focusing on regions of interest. Structural connectivity was employed to predict behavioral measures altered by AUD. This was done by applying the connectome-based predictive model (CPM) developed by Shen et al. [3], to quantify brain-behavior relationships from brain connectivity data. While CPM was initially designed for functional data, we have validated its application on structural connectivity data acquired with diffusion tensor imaging.

This study uses MRI and behavioral data from AUD and control subjects. We aim to identify biomarkers predicting the evolution of behaviors characteristically observed in AUD, such as depression, obsessive-compulsive disorders, state and trait anxiety, as well as multidimensional fatigue. The following steps were taken: Connectivity matrices were computed from tractograms obtained from MRI scans. They were afterwards inputted into CPM, which returns the most significant connections predicting the different behaviors. Biomarkers are then extracted from those connections.

Part I

State of the art

Chapter 1

Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is a medical imaging technique used to obtain structural images of the body and functional images of the brain. It respectively enables the study of the body structure and of the brain activity. It is therefore a useful and powerful tool to investigate the impact of alcohol on the brain [4].

1.1 Imaging through Nuclear Magnetic Resonance

Nuclear magnetic resonance (NMR) is the phenomenon at the core of MRI. It uses the presence of hydrogen atoms in the human body and measures the interaction of nuclear spins when placed in magnetic fields.

The human body is mainly composed of water molecules, which contain hydrogen atoms. Their nucleus are made of a proton and an electron. The proton is charged positively while the electron is charged negatively. Each proton spins around an axis $\vec{S}$ going through its center, which induces a kinetic momentum or "spin". Furthermore, when a charged particle spins, it creates a magnetic field called the magnetic moment, $\vec{\mu}$. The latter is aligned with the axis of rotation of the particle (see Fig. 1.1) [4, 5].

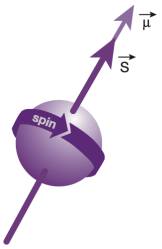


Figure 1.1: Spinning particle inducing a kinetic momentum (or spin) and a magnetic moment, $\vec{\mu}$ aligned with the axis of rotation, $\vec{S}$. From [5].

NMR measures how the nucleus' magnetization is affected when those water molecules are submitted to two distinct magnetic fields.

1.1.1 Magnetic resonance and Radio Frequency pulses

Naturally (i.e., without being submitted to any external field), the protons of those water particles are oriented at random. As a consequence, the sum of their microscopic magnetic moments is null since they cancel each other (see Fig. 1.2 left). There is thus no macroscopic magnetization vector ($\vec{M} = 0$).

However, when submitted to an external high static magnetic field, $\vec{B_0}$, protons align themselves with the direction of $\vec{B_0}$ (see Fig. 1.2 right). This suddenly creates a macroscopic magnetization vector $\vec{M}$ ($\vec{M} \neq 0$). Protons in the external magnetic field can then adopt one of the following two orientations: parallel to $\vec{B_0}$ (in the same direction as $\vec{B_0}$) or anti-parallel to $\vec{B_0}$ (in the opposite direction as $\vec{B_0}$) [4, 5].

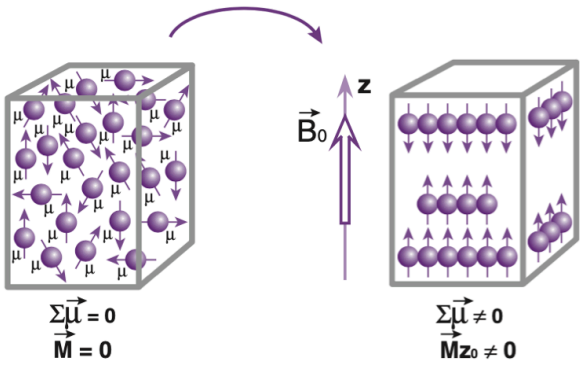


Figure 1.2: Protons without and with an external magnetic field. Adapted from [5].

Because they spin, protons are in fact not perfectly aligned to the axis of $\vec{B_0}$. They each spin with their individual angle (also called precession) and frequency around $\vec{B_0}$, as governed by the Larmor equation:

$$\omega_0 = \gamma B_0, \quad \nu_0 = \frac{\gamma}{2\pi} B_0,$$

where ω_0 is the Larmor angular frequency, ν_0 the Larmor frequency, γ the gyromagnetic factor, and B_0 the external magnetic field [4, 5, 6].

The two possible orientations a proton can adopt correspond to distinct energy levels: low for the parallel one and high in case of an anti-parallel orientation. The protons are almost equally distributed across the two energy levels. The slight surplus on the low energy level creates the macroscopic magnetization vector $\vec{M}$. $\vec{M}$ is the sum of two components: $\vec{M_z}$, the longitudinal magnetization, and $\vec{M_{xy}}$, the transverse one [4, 5, 6].

As protons follow quantum mechanics, their energy on each energy level can be quantified. A parallel proton has the energy E_1 and an anti-parallel one the energy E_2.

$$E_1 = -\gamma\hbar\frac{B_0}{2}, \tag{1.1}$$

$$E_2 = \gamma\hbar\frac{B_0}{2}, \tag{1.2}$$

where $\hbar = h/2\pi$, and h is the Planck constant. The energy difference between the two energy levels is thus equal to the following expression, derived from Eq. 1.1 and Eq. 1.2 [4, 5, 6].

$$\begin{aligned}\Delta E &= E_2 - E_1 \\ &= \gamma\hbar B_0 = \gamma\frac{h}{2\pi}B_0 \\ &= h\nu_0.\end{aligned}$$

At resting state, $\vec{M}$ is aligned to $\vec{B}_0$ and to the z-axis of the system. The longitudinal magnetization is therefore non null due to the excess protons, while the transverse components of the protons compensate each other. However, $\vec{M}$ cannot be measured while aligned with $\vec{B}_0$ due to its small size compared to the latter. In order to measure it, $\vec{M}$ must be tilted in the $x\mathcal{O}y$ plane.

This is accomplished by adding a second, rotating, magnetic field $\vec{B}_1$ applied to $x\mathcal{O}y$ along $\mathcal{O}x$. In order to change the energy level of the protons and tilt $\vec{M}$, the energy brought to the system by $\vec{B}_1$ must be equal to ΔE, the energy difference between the two different energy levels. This is called the *resonance* phenomenon. To do so, $\vec{B}_1$ must be a Radio Frequency (RF) wave and the following condition must be fulfilled:

$$\omega_r = \omega_0,$$

with ω_r the rotation frequency of $\vec{B}_1$. It leads to $\omega_1 = \gamma B_1$, with ω_1 the precession angular frequency of $\vec{B}_1$ [4, 5, 6].

The systems are thus in resonance and the magnetization vector $\vec{M}$ rotates around both magnetic fields, at the precession angular frequency ω_0 around $\vec{B}_0$ ($\mathcal{O}z$), and at the precession angular frequency ω_1 around $\vec{B}_1$ ($\mathcal{O}x$). In practice, the magnetic field or RF wave $\vec{B}_1$ is applied only for a few milliseconds. It is referred to as an *RF pulse* or *excitation*. Depending on the pulse duration, $\vec{M}$ tilts by 90° or 180°. After a 90° impulsion, the magnetization vector $\vec{M}$ is entirely in the $x\mathcal{O}y$ plane, while after a 180° impulsion, it is inverted (see Fig. 1.3) [4, 5, 6].

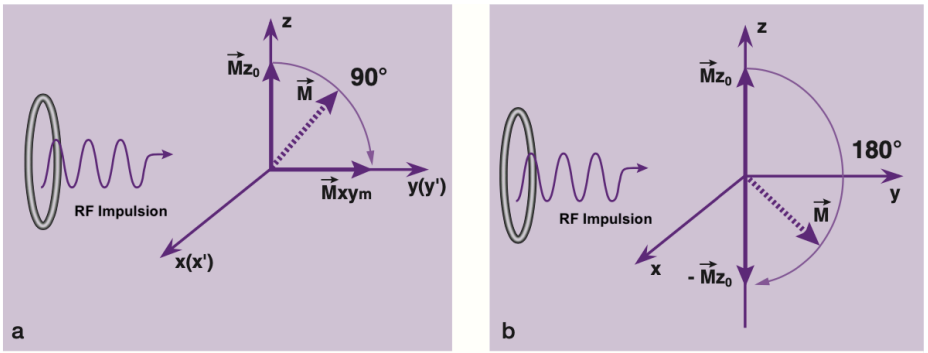


Figure 1.3: Effects of a RF impulsion on the magnetization vector $\vec{M}$: (a) 90° impulsion and (b) 180° impulsion. Adapted from [5].

During a 90° impulsion, the longitudinal component of $\vec{M}$ decreases while its transverse component increases. As soon as the impulsion stops, the magnetization vector

will go back to its initial position. The longitudinal component will thus grow back and the transverse one will disappear. This phenomenon is called the *relaxation* [4, 5, 6].

1.1.2 Relaxation

The excitation applied to the steady state leads the system to an unstable state. At the end of the RF pulse, the system returns to its original steady state, with two distinct relaxations: T_1 and T_2. T_1 is governed by the transition of protons between energy levels, and T_2 by phasing of the protons [4, 5, 6].

T_1 Relaxation

As mentioned earlier, a 90° RF pulse gives energy to the system. The protons move from the lower energy level to the higher one and the longitudinal component of $\vec{M}$ disappears. As soon as the rotating magnetic field $\vec{B}_1$ stops, the system goes back to its initial state: the protons release the energy they absorbed and the longitudinal component grows progressively. This is the *longitudinal relaxation* or *spin-lattice relaxation*. It is also called T_1 relaxation as the increase of the longitudinal component follows an increasing exponential curve with a T_1 time constant:

$$\vec{M}_z(t) = \vec{M}_0\Big(1 - e^{\frac{-t}{T_1}}\Big).$$

T_1 equals 63% of the regrowth of the longitudinal component. At the time $2T_1$, 87% has regrown, and at the time $3T_1$, the regrowth has almost finished and reached 95%. If T_1 is smaller, the longitudinal magnetization will recover quicker. Interestingly, T_1 varies as a function of the nature of the tissue [4, 5, 6].

T_2 Relaxation

In steady state, protons are not in phase. Indeed, as protons evolve in different molecular environments, small magnetic fields superimpose with $\vec{B}_0$. This creates inhomogeneities in $\vec{B}_0$, and the protons do not have the exact same Larmor frequency.

When the 90° RF pulse is applied, protons come into phase with each other and the transverse component $\vec{M}_{xy}$ appears. The protons will go out of phase as soon as the RF pulse stops. It can also be said that the protons lose their phase coherence. This leads to a rapid decay of the transverse component. It is called the *transverse relaxation* or *spin-spin relaxation*. This relaxation is also called T_2 relaxation because the decrease of the transverse component follows a decreasing exponential curve with a T_2 time constant:

$$\vec{M}_{xy}(t) = \vec{M}_0 e^{\frac{-t}{T_2}}.$$

T_2 equals 63% of the transverse component decrease. At the time $2T_2$, 87% has disappeared, and at the time $3T_2$, the disappearance has almost finished and only 5% of the transverse component remains. The transverse magnetization will decrease faster if T_2 is smaller. The time constant T_2 also varies as a function of the nature of the tissue [4, 5, 6].

Measuring NMR signals

The MRI technology measures NMR signals thanks to the *Free Induction Decay*, or *FID*. The FID is the decay of the transverse magnetization. The regrowth of the longitudinal magnetization, driven by T_1, takes more time than the decrease of the transverse magnetization, driven by T_2, which causes the magnetization vector $\vec{M}$ to rotate in a spiral. The FID stems from this rotation [4, 5, 6].

T_2^* and inhomogeneities in the magnetic field

If the magnetic field $\vec{B}_0$ was homogeneous, the measured FID signal would show an exponential decay of time constant T_2 and be fully determined by the inhomogeneities from the molecular properties of the various tissues. Unfortunately, hardware limitations also introduce constant inhomogeneities that affect the external field $\vec{B}_0$ itself. Inhomogeneities from both sources combine in the FID, which in practice follows a quicker decay, governed by T_2^* (see Fig. 1.4) [4, 5].

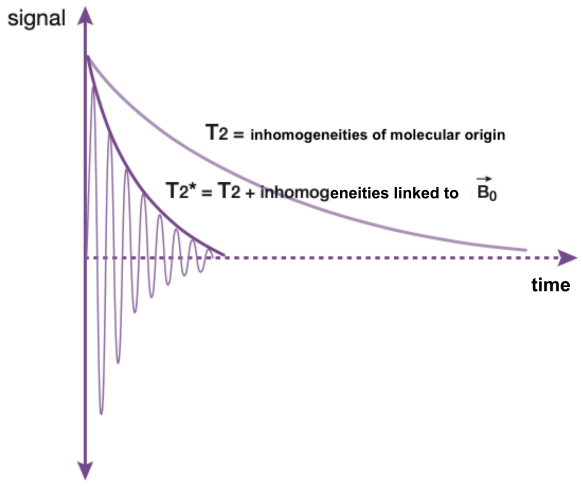


Figure 1.4: Effects of the different inhomogeneities on the time constant T_2. Adapted from [5].

1.1.3 Spin Echo sequence

The *Spin Echo (SE) sequence* is used to get rid of the hardware-driven inhomogeneities and observe T_2 instead of T_2^*. Introducing a second RF pulse of 180° cancels out the hardware-driven constant inhomogeneities linked to $\vec{B}_0$ (see Fig. 1.5) [5, 6].

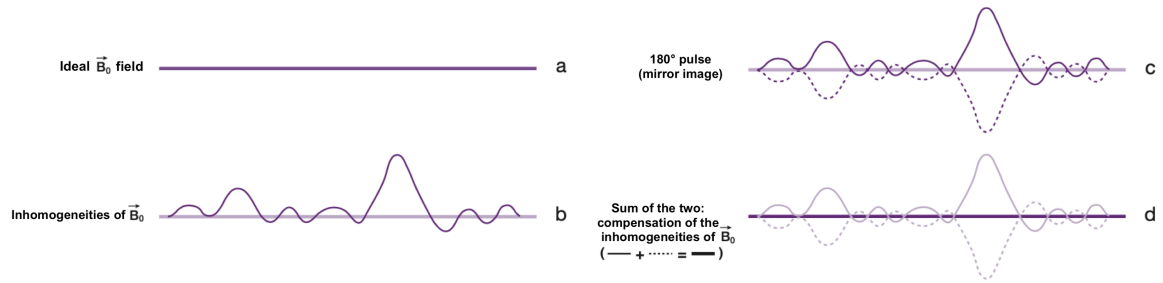


Figure 1.5: 180° *RF pulse to cancel the constant inhomogeneities linked to* $\vec{B}_0$. *Adapted from [5].*

The SE sequence is thus composed of a first RF pulse of 90°, followed by a second of 180° (see Fig. 1.6). When the 90° RF pulse is applied, the protons get in phase with each other. The spins then go out of phase due to the inhomogeneities. The 180° RF pulse is then applied. It reverses the phase shifts without changing the directions of rotation. The spins are then almost in phase again. This causes a signal called spin echo, which is then measured. The time between the 90° and the echo is called the *Echo Time (TE)*. Even then, random phase shifts persist (and cannot be suppressed). The spins not being completely in phase, the amplitude of the signal is therefore affected by the 180° RF pulse [5, 6].

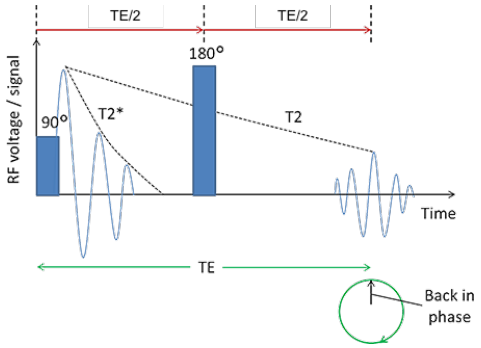


Figure 1.6: Spin Echo sequence. Adapted from [7]

A Spin Echo sequence will only capture one slice of the whole volume to image. Sequences are repeated to reconstruct the entire image. The *Repetition Time (TR)* is the time between two 90° pulses, and corresponds to the time required for the longitudinal magnetization recovery. The images obtained using SE sequences demonstrate high quality. Despite its long acquisition time, it is therefore the most used sequence in imaging techniques [5, 6].

1.1.4 Image contrast

Contrasts are the visual translation of NMR signals into different shades of grey. As explained above, measured signals capture T_1, T_2 and the proton density of the tissues. Since different tissues have different relaxation times and proton densities, they will have different shades. For example, the cerebral white matter has the shortest T_1 and T_2, the cerebrospinal fluid has the longest ones, and the grey matter has intermediate

values. It is to be noted that the relative contribution of T_1, T_2 and the proton density to the measured signal varies depending on the specific parameters of the SE sequence used for the acquisition. TR and TE values of the SE sequence thus determine the dimension (i.e. the combination of factors) along which the images are contrasted [4, 5].

T_1-weighted contrast

TR has an impact on the weight of T_1 in the contrast. A short TR maximizes the differences between tissues that have different T_1 values. Indeed, with a long enough TR, the majority of the tissues reach their initial state of longitudinal magnetization and less differences are observed. For a shorter TR, the tissues with a small T_1 will regain quicker their longitudinal magnetization and thus have a higher signal (see Fig. 1.7). They will appear in white on the image. The tissues with a high T_1 will not reach their initial state and will appear in black [4, 5, 6].

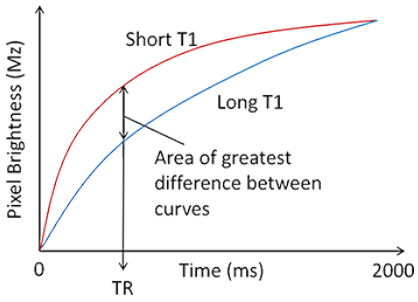


Figure 1.7: Influence of TR on the T_1-weighted contrast for 2 tissues, one with a short T_1, and the other with a long T_1. From [8].

T_2-weighted contrast

Long TR and TE are used to emphasize the contribution of T_2. The long TR minimizes the impact of T_1. TE corresponds to the time at which the signal is measured. It therefore represents the time left for the signal to decay in T_2. A longer TE enables a more accurate depiction of the variations in decay between tissues: a tissue with a long T_2 will decay slower than a tissue with a short T_2 (see Fig. 1.8). The first tissue will thus have a higher signal and will be seen in white on the image, while the second one will be black [4, 5, 6].

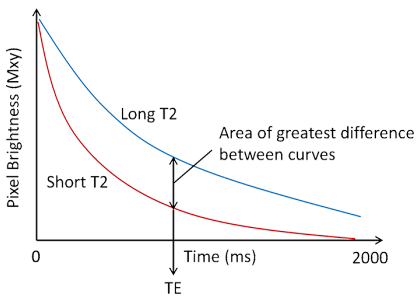


Figure 1.8: Influence of TE on the T_2-weighted contrast, for 2 tissues, one with a short T_2, and the other with a long T_2. From [8].

Proton density weighted contrast

Such a contrast is obtained with a long TR, to minimize T_1 effects, and a short TE, to minimize T_2 effects. This combination allows the longitudinal component to regrow to its initial state, and minimizes the decay of the transverse component (see Fig. 1.9). The only difference that can be seen between the tissues is then a difference of proton density [4, 5].

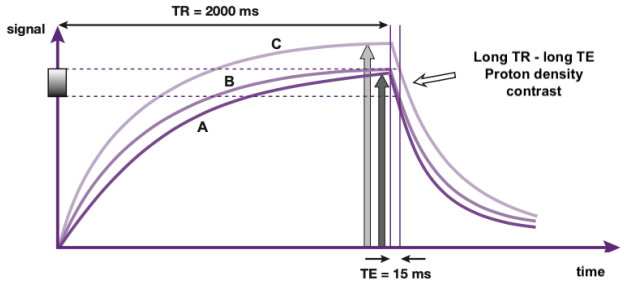


Figure 1.9: Proton density weighted contrast obtained by choosing a long TR and a short TE in order to only observe the variation in the amplitudes of the signal. Adapted from [5].

1.1.5 Acquisition

In order to reconstruct the image, spatial information must be added to the different signals. It is done by adding three magnetic field gradients, one for each axis. This is called *encoding*. A *Slice Selection Gradient* (GSS) is used to select the anatomical volume of interest (z-axis). Afterwards, a *Phase Encoding Gradient* (GPE) and a *Frequency Encoding Gradient* (GFE) are applied to encode the vertical (y-axis) and horizontal (x-axis) positions within this volume. All these gradients have identical properties but are applied in different directions and in sequence [5, 6, 9].

First, the slice plane is selected by applying the slice selection gradient, perpendicular to the desired slice plane. The GSS is applied on top of B_0, and the protons are affected by it. Their resonance frequency then varies proportionately to the GSS, according to the Larmor equation. A RF wave, the *selective RF pulse*, is applied at the same moment. It has the same frequency as the protons in the targeted slice plane. Only the protons in this plane will thus be affected and will emit a signal [5, 6, 9].

The second step applies a phase encoding gradient, which provides the vertical position. The GPE modifies the spin resonance frequencies during its application, and induces a persisting dephasing. The protons precess all at the same frequency but in different phases. The protons in the same row, perpendicular to the direction of the gradient, are all having the same phase. This phase difference persists up to when the signal is recorded [5, 6, 9].

The final step applies a frequency encoding gradient, to provide the horizontal position. As long as the GFE is applied, all protons in a column have an identical Larmor frequency [5, 6, 9].

1.2 Diffusion Weight Imaging

Diffusion Weight Imaging (DWI) uses the movements of the water molecules inside the brain to characterize its microstructures. Those movements are Brownian and are constrained by the tissues and structures of the brain. The diffusivity of a water molecule is indeed directly impacted by the structure of its surroundings. It is therefore used to extract information on the structure of the tissues in a non-invasive way. Two types of movements can occur, *free diffusion* and *restricted diffusion* [5, 9]:

- **Free Diffusion:** Water molecules can move freely in every direction of the space without any constraint. Such a motion happens in the cerebrospinal fluid for example [9].
- **Restricted Diffusion:** The movements of the water molecules are restricted. There are two different kinds of restrictions. Restricted *isotropic* diffusion defines a situation where the motions are restricted in the same way for all the directions in space. In a restricted *anisotropic* diffusion case, the motion is oriented in a certain direction. There are thus restrictions of the diffusion in only some of the directions. This case occurs in nerve fibers for example [9].

Diffusion imaging can therefore give information on the microstructure properties of the tissue around the water molecules. The information given by the motion of the water particles is an integral part of the MRI signal and can be obtained. The signal is affected solely by the molecular displacements oriented in the direction of the gradient, and thus visible. A new sequence of higher strength must be used to highlight the diffusion effects, the *Pulsed Gradient Spin-Echo sequence* (see Fig. 1.10). It is composed of 2 gradients, one applied before and one after the 180° RF pulse. For immobile water molecules, the first diffusion gradient induces a phase shift, which is perfectly balanced by the second gradient. For moving water particles, the second gradient will not rephase the spins. As water molecules diffuse more rapidly, they become more dephased and the recorded signal becomes weaker [5, 9].

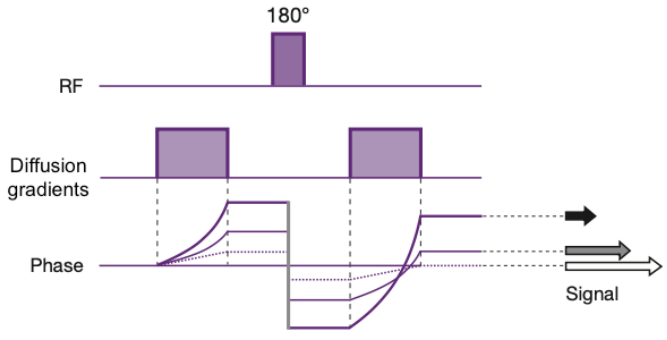


Figure 1.10: Pulsed Gradient Spin-Echo sequence and effects on the phase of the water molecules. The different arrows represent the attenuation of the signal depending on the motion of the water particles. the white arrow corresponds to immobile water molecules while the grey arrow represents slow moving molecules, and the black arrow fast moving molecules. Adapted from [5].

Changing the amplitude and application time of the diffusion gradient can improve the attenuation of the MRI signal. The *b-factor* is determined by those two characteristics. It quantifies the applied diffusion weighting of a pulse sequence in a single number.

$$b = \gamma^2 G^2 \tau^2 \left(\Delta - \frac{\tau}{3} \right) \quad [s/mm^2],$$

where γ is the gyromagnetic ratio, G the gradient amplitude, τ the gradient application time, and Δ the interval between the centers of the two diffusion gradients. Increasing the value of b increases the sensitivity of the sequence to molecular diffusion. A high b-value scan will exhibit more signal attenuation of diffusion [5, 10].

1.2.1 Diffusion Tension Imaging (DTI)

Diffusion Tensor Imaging is used to characterize the diffusion, its anisotropic coefficient, preferred directions and restrictions in space. It is used to analyze the 3D shape of the diffusion thanks to a tensor. It fully describes the motion of the molecules in each direction and the correlation between these directions. It is defined as a 3 by 3 symmetric matrix [9, 10, 11, 12].

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

A minimum of six Diffusion-Weighted sequences are needed to acquire the tensor since $D_{xy} = D_{yx}$, $D_{xz} = D_{zx}$, and $D_{yz} = D_{zy}$. A seventh measure with a low b-factor is also performed to obtain a baseline image. Different metrics can be computed thanks to λ_1, λ_2, and λ_3, the three eigenvalues of the tensor [9, 10, 11].

Mean Diffusivity (MD)

Mean diffusivity is the estimation of the overall diffusion in a *voxel*, i.e., the 3D equivalent of a pixel. It also characterizes the amount of water in the extracellular space [10, 12].

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

Axial Diffusivity (AD)

Axial diffusivity is the diffusivity in the main axis of the anisotropic diffusion ellipsoid. It can be interpreted as the diffusivity parallel to a white matter fiber tract, and will thus make the most sense in voxels with coherently oriented axons and no fiber crossings. A low value may denote a less coherent orientation of axons, axonal injury, or reduced axonal caliber [10, 13].

$$AD = \lambda_1.$$

Radial Diffusivity (RD)

Radial diffusivity characterizes the diffusivity perpendicular to the principal axis of diffusion. As for the AD, this metric makes the most sense in voxels with coherently oriented fibers and no fiber crossing. A low value may indicate myelin perturbations [10, 13].

$$RD = \frac{\lambda_2 + \lambda_3}{2}.$$

Fractional Anisotropy (FA)

Fractional Anisotropy is the ratio between the anisotropic component of D and the whole diffusion tensor. It indicates the shape of the diffusion tensor, and the degree of anisotropy of the voxel (see Fig. 1.11).

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

The value ranges from 0 to 1, 0 meaning that the tensor is completely isotropic, and 1 completely anisotropic [10, 11].

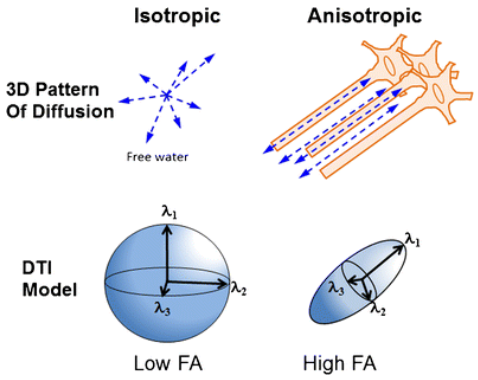


Figure 1.11: DTI Tensor depending on the type of diffusion. From [14]

1.2.2 Tractography

Tractography allows to estimate trajectories of the white matter fibers using diffusion tensors. It is based on local estimates of the diffusion direction. Once we know the fibers orientation at each voxel, these directions can be linked to reconstruct complete tracks. This basically involves starting from a seed location and following the preferred direction until we reach a new voxel. We can then switch to the new voxel preferred direction and continue until an entire track is spread [11, 15, 16].

There are two main families of tractography algorithms: *deterministic algorithms* and *probabilistic algorithms* (see Fig. 1.12). In both algorithms, the first step is the choice of a seed point.

In deterministic algorithms, a tract is then generated according to the diffusion orientation in the seed voxel. The tract direction is continuously updated according

to the new voxel orientation. As a result, a tract will be found. Tracking is ended once it reaches a voxel unlikely to belong to the track. To determine when the tract ends, a curvature and a FA threshold are defined, such that a track can not continue if its curvature is too sharp or if it reaches a voxel where the FA is too low. However, the diffusion orientation estimate may contain uncertainties. Tractography will thus be impacted by those, and the uncertainties will propagate or accumulate as the track goes. This results in tracts terminating prematurely or the target pathway not being found [11, 16].

In probabilistic algorithms, uncertainties of the orientation estimates are taken into account. A distribution of fiber orientation is estimated for each voxel. The tracking direction at each point along the path is chosen at random from the distribution of possible directions. A tract probability map can be obtained. Such a method is computationally much more demanding [11, 16, 17].

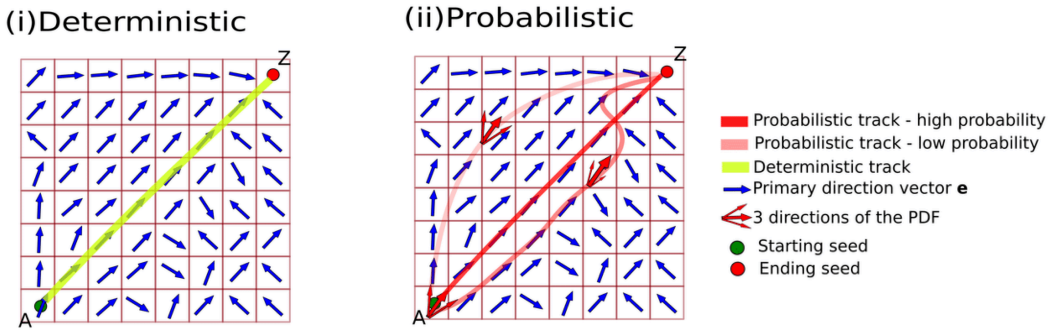


Figure 1.12: Example of deterministic tractography (i) and probabilistic tractography (ii). The track obtained from the deterministic approach is in yellow, while the tracks found by the probabilistic approach are in a scale of red, depending on their probability of existence. Adapted from [15].

Since tractography is later used in this thesis to get connectivity measures, the parameters used to produce it affect the results obtained. These parameters are the number of streamlines used, and the different thresholds defined to determine the end of the different tracts. The number of streamlines impacts the weights of the connectivity matrices (more details on connectivity matrices are available in the section 2.2). The end thresholds of the tracts will of course affect where they stop, and therefore the number of tracts between each pair of regions. They thus also affect the connectivity matrices.

Chapter 2

Graph Theory and Connectivity

The aim of this section is to explain how graph theory can be used to represent the connectivity of the brain.

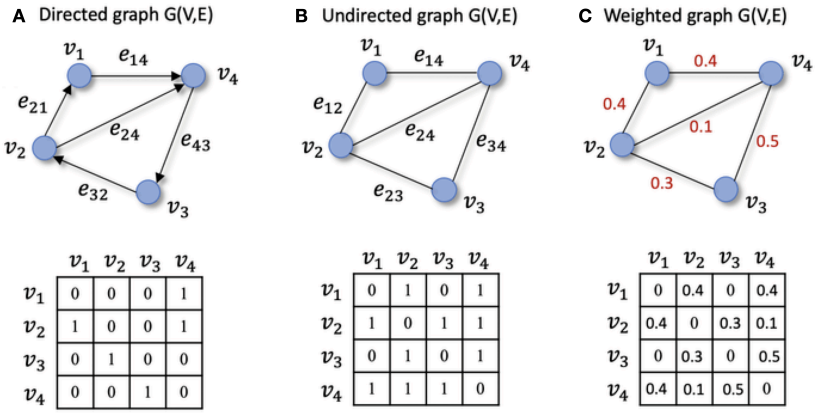
2.1 Graph Theory

Graphs are employed to model the relationships between different objects using nodes and edges. The brain is a highly connected organ and can be represented by a graph. In this context, the nodes correspond to the brain regions or voxels, while edges represent the links and the connectivity between those regions [18, 19, 20].

Graphs have different properties. They can either be directed or undirected, as well as weighted or unweighted (binary). In a directed graph, edges are directed from a specific node to another. In an undirected graph, edges represent links between different nodes, without any indication of direction. All the edges will have an equal weight in a binary graph. In a weighted graph, every edge will have an assigned weight. Some edges will thus be more important than others [18, 19, 20].

2.2 Connectivity

Graphs can be represented as matrices. One possible representation of a graph is the adjacency matrix. It is a $N \times N$ matrix, where N is the number of nodes. The rows and columns represent the nodes of the graph while the matrix entries contain the edges. The matrix entry between the row i and column j is the number of connections between the region i and the region j. Such a matrix can thus be used to retain information about the brain connectivity. In neuroscience, the entire map of cerebral connections is called a connectome [19, 20, 21, 22].



	v_1	v_2	v_3	v_4
v_1	0	0	0	1
v_2	1	0	0	1
v_3	0	1	0	0
v_4	0	0	1	0

	v_1	v_2	v_3	v_4
v_1	0	1	0	1
v_2	1	0	1	1
v_3	0	1	0	1
v_4	1	1	1	0

	v_1	v_2	v_3	v_4
v_1	0	0.4	0	0.4
v_2	0.4	0	0.3	0.1
v_3	0	0.3	0	0.5
v_4	0.4	0.1	0.5	0

Figure 2.1: Examples of graphs with their adjacency matrix: (a) directed and binary graph, (b) undirected and binary graph, (c) undirected and weighted graph. Adapted from [23].

In the context of this thesis, structural connectivity matrices are used. They contain an estimate of the connection strength among all regions. They are obtained by processing diffusion-weighted MRI data to produce a tractogram. It then provides information on the number of connections between brain regions. This results in a weighted connectivity matrix [24].

Chapter 3

Alcohol Use Disorder

3.1 Definition

Alcohol Use Disorder (AUD) is "a medical condition characterized by an impaired ability to stop or control alcohol use despite adverse social, occupational, or health consequences" [1].

Alcohol consumption has a huge impact on the world population nowadays. In many countries, drinking is a part of the daily life, and AUD affects 1.4% of the global population. Alcohol consumption is not equally spread around the world (see Fig. 3.1). In most European countries, more than two-thirds of the adult population drink alcohol, meanwhile the consumption is significantly lower in North Africa and the Middle East. This disorder is not affecting only adults. 414,000 children aged 12 to 17 have AUD. It corresponds to 1.7% of this age group. However, the age group with the highest prevalence of AUD is between 25 and 34 [1].

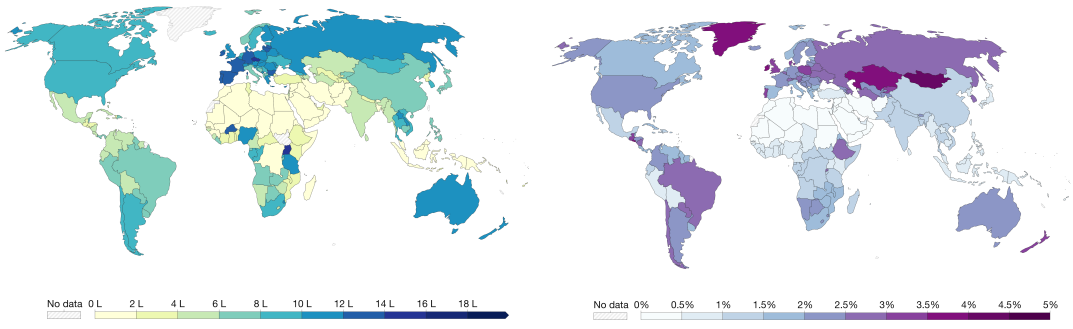


Figure 3.1: Worldwide situation in terms of alcohol consumption in liters of pure alcohol per person aged 15 or more in 2018 (left), and share of population with an alcohol use disorder in 2019 (right). Adapted from [25].

A diagnosis of AUD is made if at least two of the following symptoms are observed within the last year. The severity of the Alcohol Use Disorder will vary according to the number of symptoms experienced [1].

- Intention and time: Drinking more or longer than intended, spending more time drinking or recovering,

- Craving and failed attempts to cut down alcohol consumption,
- Interference between the use of alcohol and the care of their home, family, work or school,
- Social and lifestyle consequences: Drinking even though it caused trouble with family or friends, giving up or cutting back on hobbies or pleasures in order to drink,
- Risks and psychological consequences: Being more than once in dangerous situation (driving, swimming, ...) because of alcohol, drinking despite feeling anxious, depressed, or having memory blackouts,
- Building a tolerance to the effects of alcohol and having to drink more in order to feel less or the same effects as before,
- Having withdrawal syndromes such as trouble sleeping, shakiness, restlessness, nausea, sweating, a racing heart, or a seizure.

Even though alcohol use is common across the globe, it remains a risk factor for premature deaths. In fact, it causes three million premature deaths per year, 5.3% of all deaths worldwide, making it the seventh deadliest risk factor in 2019. Alcohol dependence also causes indirect deaths. In fact, a person with an alcohol dependence is ten times more likely to commit suicide [2, 25].

3.2 Impact of AUD

Despite the well-known negative effects of excessive alcohol consumption on health, it still remains one of the most commonly used drugs that leads to brain damage. Alcohol is known to be neurotoxic and directly impacts nerve cells. Changes in brain structure and function, as well as a loss of brain mass are observed in brains impacted by AUD. These changes may contribute to the development and progression of the addiction. AUD has multiple impacts on the brain, which can be divided into two categories: biological impacts on the brain structure and functional changes affecting the behavior [26].

3.2.1 Biological impact

Alcohol Use Disorder impacts the structure of the brain. Structural and volumetric changes can be observed because of a significant alcohol consumption. MRI studies showed a loss of volume and a decreased myelination throughout the brain. Differences in grey and white matter are observed. AUD is specifically associated with a loss of grey matter volume. It particularly affects regions such as the prefrontal cortex, cingulate cortex, insula, and striatum. The amount of grey matter volume lost is linked to lifetime alcohol consumption and duration of alcohol dependence. Differences in white matter volume and microstructure can also be seen, especially in the corpus callosum. It indicates potential disruption of myelination and axonal integrity. The effects of alcohol on the brain are not uniform, with certain brain areas or cell populations

being more vulnerable. These regions include the prefrontal cortex, hippocampus, cerebellum, white matter, and glial cells [26, 27].

3.2.2 Behavioral and psychological impact

Alcohol does not only alter the structure of the brain, but also affects the functional regulation of important brain systems controlling various behaviors such as reward processing, impulse control, and emotional regulation. Alcohol intoxication can also increase risky behaviors, and AUD affects memory, cognition, motivation, and motor capabilities. AUD is often associated with psychiatric disorders, such as anxiety, mood, and depressive disorders. The latter is the most common among individuals with Alcohol Use Disorder. AUD and depression reinforce each other since depressive disorders increase the risk for AUD, and AUD increases the risk for depressive disorders [27, 28, 29].

3.3 Abstinence and relapse

For the majority of individuals suffering from alcohol dependence, abstinence reverses partially the damage done by AUD on the brain. If started in time, it can effectively counteract a significant amount of the damage caused by excessive drinking. Effects can already be visible within two weeks, such as the beginning of the regeneration of the lost grey matter volume. After one month of abstinence, an increase in volume of the cortical grey matter, overall brain tissue, and hippocampal tissue can be observed. After six months of abstinence, the brain continues to regenerate itself, and cognitive improvements are visible. People who maintain sobriety over the long term show a decrease of the size of the third ventricle or a global increase of the brain volume, especially in the temporal and frontal regions. Inversely, individuals who relapse show an increase of the size of the third ventricle and a reduction of white matter or overall brain tissue. Some regions are more responsive to abstinence or relapse. For example, the volume of the cortical white matter is particularly responsive to recovery with sustained abstinence, and vulnerable to relapse. Prolonged sobriety improves or stabilizes measures of brain tissue volume while relapsing reverses the rapid regeneration occurring soon after abstinence [27, 28, 30, 31].

Maintaining abstinence is thus the optimal outcome for most AUD patients. The risk of relapse decreases considerably after 10 to 14 weeks of abstinence. Unfortunately, it is not often reached, with over 80% of alcohol dependent patients relapsing within a year despite the current available therapies. Even though neuroimaging studies have helped our understanding of the impact of alcohol on the brain, this has not led so far to huge improvements in treatment efficacy. The identification of biomarkers should help to develop more effective therapeutic targets [27, 28, 30, 31].

3.4 Analysis of connectivity and AUD

Many studies about disruption of functional networks in individuals with AUD were conducted [32], but few about structural changes in connectivity. This section

aims to present the findings regarding altered functional connectivity within different networks, such as the *Default Mode Network*, the *Executive Control Network*, the *Salience Network* and the *Reward Network*. These results will later be compared with those of this thesis, to see whether changes in functional connectivity also translate into structural connectivity alterations. Table 3.1, in section 3.4.9, summarizes the results found in the literature.

3.4.1 Default Mode Network (DMN)

The Default Mode Network is a set of brain regions that show a high brain activity at rest. It includes the posterior cingulate cortex (PCC), retrosplenial cortex, inferior parietal cortex, medial prefrontal cortex, medial temporal lobe, and precuneus (see Fig. 3.2). It was found to be affected by AUD, and disruptions within it may affect behavior [32, 33, 34].

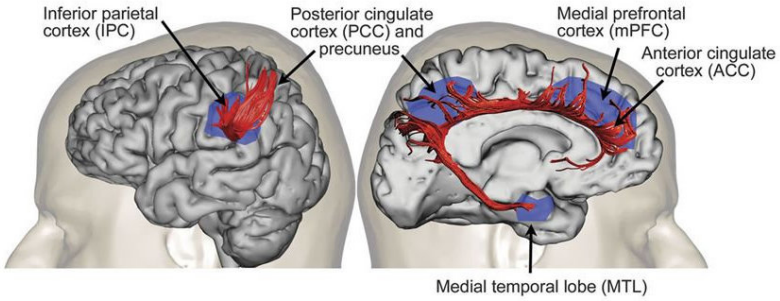


Figure 3.2: Default Mode Network. Adapted from [35].

When compared to control participants, individuals suffering from alcohol dependence display an expanded PCC connectivity with the hippocampus, the middle temporal gyrus and pole, while having a smaller PCC connectivity with the right medial frontal gyrus. A significantly greater PCC connectivity to VI cortices was found in AUD patients and weaker connectivity to dorsal striatal regions counting the caudate nucleus and portions of the anterior cingulate cortex. Moreover, AUD patients showed compromised functional connectivity between the posterior cingulate cortex and cerebellar regions compared to controls. However, connectivity between the left posterior cingulate cortex and left cerebellar regions was stronger for AUD patients. Such reinforced connections suggest the presence of compensatory networks. These differences in connectivity were not observed for the right posterior cingulate cortex [32, 33, 34].

3.4.2 Executive Control Network (ECN)

The Executive Control Network includes the prefrontal, the frontopolar, the anterior cingulate, the posterior parietal cortices, the cuneus, the basal ganglia, and the limbic system. It plays a role in goal-directed behavior, emotion processing, learning, memory, habits, and attention. AUD damages particularly this network, and the alterations caused seem to impair multiple cognitive, emotional, and sensorimotor processes [32, 33, 36, 37].

AUD patients showed a stronger connectivity than controls between the superior frontal gyrus and the cuneus, the fusiform, and VI association cortices, the right superior temporal gyrus, and the medial frontal gyrus. They also showed a decreased connectivity between the superior frontal gyrus and the middle cingulate cortex, the orbitofrontal regions, and the inferior parietal lobe. The weakened connection between the superior frontal gyrus and a right middle cingulate cortex region extending into the caudate body correlates with more depressive symptoms and higher anxiety [32, 33, 38].

AUD affects especially the left executive control network. The connectivity strength of the left executive control network for AUD patients correlates negatively with failed control of alcohol consumption, alcohol disorder severity, and measures of hazardous drinking. The connections between the parietal regions and the dorsolateral prefrontal, middle frontal, and temporal regions in the left executive control network drive these relationships. The frontal lobe and limbic systems, basal forebrain, and cerebellum also play an important role in cognitive processes and the inability to control alcohol consumption. AUD patients show a reduced connectivity strength between the nucleus accumbens and thalamus, caudate, postcentral, and parietal regions [32, 33, 38].

3.4.3 Salience Network (SN)

The salience network includes the anterior insula, anterior cingulate cortex, and ventral striatum. It also includes nodes in the amygdala, hypothalamus, and thalamus. It plays a role in the interoception of the feelings linked with reward. It is also implicated in both drug craving and the relapse that often occurs in addiction [39, 40].

A stronger connectivity was found in AUD patients between the anterior cingulate cortex and bilateral frontostriatal regions such as the caudate head and inferior prefrontal, orbitofrontal, dorsal anterior cingulate cortices. AUD patients displayed a weaker connectivity between the anterior cingulate cortex and bilateral inferior parietal lobe, middle cingulate cortex, posterior cingulate cortex, precuneus, and thalamus [32].

3.4.4 Attention network (AN)

The attention network includes the intraparietal sulcus and frontal eye fields. It plays a role in visuospatial attention [40, 41].

In AUD patients, the superior parietal lobe has a stronger connectivity with the left precentral and middle frontal gyri, bilateral postcentral and supramarginal gyri, and temporal cortex regions, and a weaker connectivity with the left inferior frontal and orbitofrontal gyri, and with cerebellar regions. A stronger connectivity between the superior parietal lobe and the left middle frontal gyrus correlated with less depressive symptoms for AUD patients [32].

3.4.5 Somatosensory network (SsN)

The somatosensory network processes sensory information such as temperature, touch, proprioception, or taste. It interprets different signals coming from various parts of the body [42, 43].

AUD patients have a stronger connectivity between the postcentral gyrus and the bilateral parietal lobes, the left cuneus, the superior and middle occipital gyrus, the right lingual gyrus, and the cerebellum. It also has a weakened connectivity with the left temporoparietal junction and the left inferior temporal gyrus [32].

3.4.6 Reward network (RN)

AUD patients show expanded connectivity of the Nucleus Accumbens (NAcc) with the bilateral medial and inferior prefrontal gyri, as well as the left inferior temporal gyrus. They also exhibit restricted connectivity of the NAcc with the bilateral amygdala, the left medial temporal, the angular, and the middle occipital gyri. Moreover, expanded connectivity between the NAcc and the medial prefrontal gyrus correlates positively with higher scores on trait anxiety. Furthermore, restricted connectivity between the NAcc and the left amygdala correlates with higher levels of alcohol craving, and lower levels of self-esteem. Additionally, weaker connectivity between the NAcc and the left inferior parietal lobule is linked with more depressive symptoms [32].

3.4.7 Cerebellum

The cerebellum has different functions, such as maintaining balance, coordinating movement, and motor learning [44].

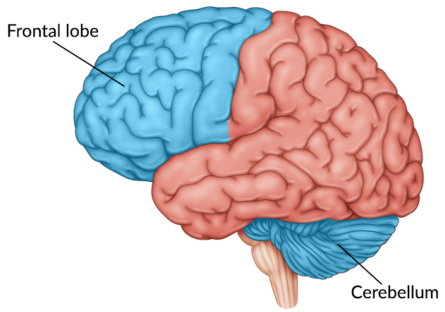


Figure 3.3: Cerebellum and frontal lobe. Adapted from [45] and [46].

It plays a role in compensating for functional damages of the frontal regions. It is a particularly vulnerable brain region. Functional connectivity between the frontal cortex (specifically Brodmann Area 9) and the Lobule VIII of the inferior cerebellum is weaker in AUD patients than controls (see Fig. 3.3). A reduced connectivity is also exhibited between the premotor areas (Brodmann Area 6) and the Lobule VI of the superior cerebellum in AUD patients [33].

3.4.8 Other neural links affected by AUD

AUD also impacts the corpus callosum. It is a bundle of white matter fibers establishing connections between cortical regions across both brain hemispheres. Impairments can lead to difficulties with the integration of sensory and cognitive information. A disruption of commissural connectivity within the corpus callosum was observed. The genu (anterior region) and splenium (posterior region) are particularly affected [47, 48].

AUD patients exhibit a weakened connection between the amygdala and the globus pallidus (pallidum and putamen), which may reduce the emotionally driven motivational outputs. The connections between the left insula and the anterior cingulate cortex, the ventrolateral prefrontal cortex, and the medial orbitofrontal cortex are also weakened [49].

AUD patients that relapsed several times show an expanded connectivity between the left insula and the right and left colliculus neuronal clusters. It is also the case between the amygdala and the Bed Nucleus of Stria Terminalis. The connectivity strength between the amygdala and the dorsolateral prefrontal cortex correlates positively with the Impaired Control Questionnaire (ICQ) score. However, the ICQ score correlates negatively with the connectivity strength between the left insula and the inferior frontal cortex, as well as between the insula and the frontal pole on both sides. This connection also correlates negatively with the Severity of Alcohol Dependency Questionnaire (SADQ) score [33, 49].

Finally, AUD patients show altered connectivity between the bilateral ventral striatum and different regions of the prefrontal cortex, such as the lateral orbitofrontal cortex, the medial prefrontal cortex, and the dorsolateral prefrontal cortex. The altered connectivity between the ventral striatum and dorsolateral prefrontal cortex may play a role in shifting from goal-directed to habitual actions. It may thus control compulsive alcohol consumption. Furthermore, the connectivity between the nucleus accumbens and the medial prefrontal cortex is associated with drinking characteristics (duration of alcohol consumption, the number of drinks per use, frequency and severity of drinking, and magnitude of alcohol craving) [50].

3.4.9 Summary

Table 3.1: Impact of AUD on the connectivity strength of different neural links, as well as the links with behavioral changes due to AUD, and the belonging of the neural links to a specific neural network if so.

Neural links	**Impact of AUD on connectivity**	**Link with behaviors**
Default mode network		
PCC with the hippocampus, the middle temporal gyrus and pole	Expanded	
PCC with right medial frontal gyrus	Weakened	
PCC with VI cortices	Expanded	
PCC with dorsal striatal regions (caudate nucleus and portions of the anterior cingulate cortex)	Weakened	
PCC with cerebellar regions	Weakened	
Left PCC with left cerebellar regions	Expanded	
Executive control network		
Superior frontal gyrus (SFG) with the cuneus, the fusiform, and VI association cortices, the right superior temporal gyrus, and the medial frontal gyrus	Expanded	
SFG with middle cingulate cortex (MCC), the orbitofrontal regions, and the inferior parietal lobe	Weakened	
SFG with right MCC extending into the caudate body	Weakened	Positive correlation with more depressive symptoms and higher anxiety
Parietal regions with the dorsolateral prefrontal, middle frontal, and temporal regions		Negative correlation with failed control of alcohol consumption, alcohol disorder severity, and measures of hazardous drinking
NAcc with thalamus, caudate, postcentral, and parietal regions	Weakened	
Salience network		

Anterior cingulate cortex (ACC) with bilateral frontostriatal regions (caudate head and inferior prefrontal, orbitofrontal, dorsal anterior cingulate cortices)	Expanded	
ACC with bilateral inferior parietal lobe, middle cingulate cortex, posterior cingulate cortex, precuneus, and thalamus	Weakened	
Attention network		
Superior parietal lobe with left precentral and middle frontal gyri, bilateral postcentral and supramarginal gyri, and temporal cortex region	Expanded	
Superior parietal lobe with left inferior frontal and orbitofrontal gyri, and with cerebellar regions	Weakened	
Superior parietal lobe with left middle frontal gyrus	Expanded	Positive correlation with less depressive symptoms
Somatosensory network		
Postcentral gyrus with the bilateral parietal lobes, the left cuneus, the superior and middle occipital gyrus, the right lingual gyrus, and the cerebellum	Expanded	
Postcentral gyrus with the left temporoparietal junction and the left inferior temporal gyrus	Weakened	
Reward network		
NAcc with the bilateral inferior prefrontal gyrus, and the left inferior temporal gyrus	Expanded	
NAcc with the bilateral amygdala, the left medial temporal, the angular, and the middle occipital gyri	Weakened	
NAcc with the medial prefrontal gyrus	Expanded	Positive correlation with higher trait anxiety scores
NAcc with the left amygdala	Weakened	Positive correlation with higher levels of alcohol craving, lower levels of self-esteem
NAcc with left inferior parietal lobule	Weakened	Positive correlation with more depressive symptoms

Other neural links		
Frontal cortex (specifically Brodmann Area 9) with Lobule VIII of the inferior cerebellum	Weakened	
Premotor areas (Brodmann Area 6) with the Lobule VI of the superior cerebellum	Weakened	
Commissural connectivity within the corpus callosum, especially the genu and the splenium	Disrupted	
Amygdala with the globus pallidus (pallidum and putamen)	Weakened	Reduction of the emotionally driven motivational outputs
Left insula with the anterior cingulate cortex, the ventrolateral prefrontal cortex, and the medial orbitofrontal cortex	Weakened	
Left insula with the right and left colliculus neuronal clusters	Expanded	
Amygdala with the bed nucleus of stria terminalis	Expanded	
Amygdala with the dorsolateral prefrontal cortex		Positive correlation with the ICQ score
Left insula with the inferior frontal cortex		Negative correlation with the ICQ score
Insula with the frontal pole on both sides		Negative correlation with the ICQ score and with the SADQ score
Bilateral ventral striatum with regions of the prefrontal cortex, such as the lateral orbitofrontal cortex, the medial prefrontal cortex, and the dorsolateral prefrontal cortex	Altered	
Ventral striatum with the dorsolateral prefrontal cortex	Altered	Role in shifting from goal-directed to habitual actions. May control compulsive alcohol consumption

NAcc with the medial prefrontal cortex		Associated with drinking characteristics (duration of alcohol consumption, the number of drinks per use, frequency and severity of drinking, and magnitude of alcohol craving)

Part II

Analysis

Chapter 4

Methods

The goal of this thesis is to identify biomarkers for patients suffering from alcohol dependence by studying their brain connectivity. A global approach was taken, using the whole brain and not focusing on regions of interest. All the steps taken to reach the aim of this thesis are represented in Fig. 4.1.

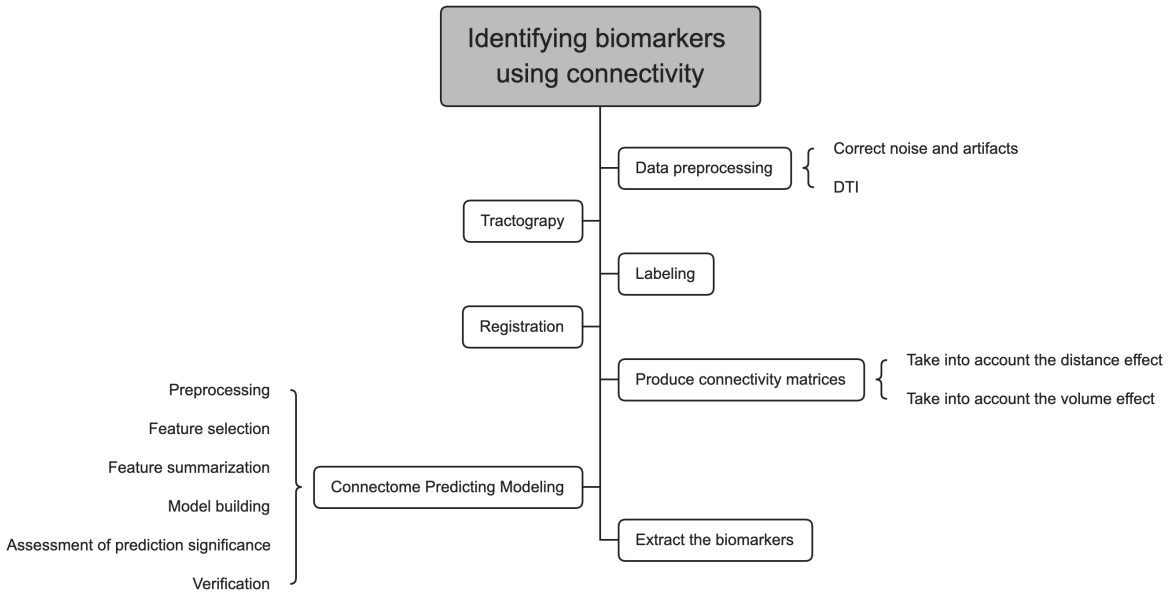


Figure 4.1: Steps to reach the aim of this thesis.

First, brain connectivity was used to predict certain behaviors altered by alcohol consumption. This step was inspired by the connectome-based prediction method proposed in the paper *"Using connectome-based predictive modeling to predict individual behavior from brain connectivity"* by Shen X. et al. [3]. Connectome-based Predictive Modeling (CPM) is a data-driven approach that uses connectivity data and employs cross-validation to build predictive models of brain-behavior connections.

The second key step is to extract the biomarkers, i.e., the connections that play the biggest roles in predicting those behaviors.

Preliminary steps were also needed. First, the data was preprocessed, and tractography was done to produce the connectivity information. Afterwards, a brain region was assigned to each voxel of the brain, and the data was registered. Finally, the

obtained connectivity matrices were normalized.

This chapter will review the different methods used, the reasons behind their choices, as well as their implementations.

4.1 Data collection

The data used in this thesis was acquired at *Cliniques universitaires Saint-Luc* for the PhD thesis of Mélissa Salavrakos, on 73 participants: 53 alcohol dependent participants, and 20 control participants. The patients suffering from AUD were carefully chosen upon a set of criteria: a highly frequent alcohol consumption before the participation in a withdrawal program, no severe comorbidity, no regular or significant intake of anti-inflammatory drugs, and no contraindication to MRI. This data was already used in the scope of another master thesis, *"Identifying the microstructural changes for alcoholic patients during short-term abstinence via diffusion Magnetic Resonance Imaging"* by Manon Dausort and Faustine Falyse, which contains a detailed description of the data collection [51].

4.1.1 Structural data

As stated before, the patients suffering from alcohol dependence participated in a withdrawal program. They performed MRI scans at different times during the course of the study. The first scan took place on the day of admission to the withdrawal program, and is further referenced as Time 1 (t_1). The second scan took place after 18 days of abstinence and is called Time 2 (t_2). A final scan was performed after 3 months of abstinence and is therefore called Time 3 (t_3). Unfortunately, some patients relapsed or did not show up for some of the t_2 or t_3 scans. They were thus excluded from the sample. Only the scans in t_1 and in t_2 were performed on control participants.

The MRI device used to collect the data is the *3T Signa^{TM} Premier* by *GE Healthcare*, with an Echo Planar Spin Echo sequence. The imaging frequency used during acquisition was 127.77 Hz, with a slice thickness of 2 mm. The TR for image acquisition was set at 4.837 seconds, while the TE varied slightly among patients but always remained around a time of 79-80 ms [51]. A total of 107 scans are usable for the participants with AUD, and 39 for the control participants.

For each patient was gathered a set of 3×2 scans for each time: two T_1-weighted contrasted anatomic scans, two diffusion scans, as well as two diffusion correction scans.

4.1.2 Behavioral data

Each of the participants (controls and patients) also completed the following behavioral questionnaires:

- *Beck Depression Inventory (BDI)*: scale to evaluate the severity of depression [52, 53]. It consists of 21 questions with four possible answers each. To each answer is

assigned a score, ranging from 0 to 3. The total score of the test is thus between 0 and 63 (= 3∗21). A higher score reflects a higher number of depressive symptoms reported by the participant. A total score between 0 and 9 indicates minimal depression, between 10 and 18 mild depression, between 19 and 29 moderate depression, and between 30 and 63 severe depression.

- *Obsessive Compulsive Drinking Scale (OCDS)*: evaluation of alcohol craving [54, 55]. It consists of 14 questions that provide a total craving score, as well as a compulsion and an obsession score.
- *State-Trait Anxiety Inventory (STAI)*: measure of trait and state anxiety for adults [56]. It is commonly used to distinguish depression from anxiety. It consists of 20 questions with each 4 answers. To each answer is assigned a score ranging from 1 to 4. The total score of the test is thus between 20 and 80 (4*20). A high score indicates a high level of anxiety.
- *Multidimensional Fatigue Inventory (MFI)*: measure of the fatigue level around 5 dimensions: general fatigue, physical fatigue, reduced motivation, reduced activity, and mental fatigue [57]. It consists of 20 questions.

The protocol included the collection of additional behavioral and biographical information not used in the present thesis, such as the Brief evaluation of alcohol-related neuropsychological impairment (BEARNI) test, the Cushman score and personality measures.

Some patients did not complete all the tests. 19 of the 20 control participants did the behavioral tests. For AUD participants, 51 behavioral measures were gathered in t_1 for each behavior test (expect for MFI which has only 47 measures), 48 in t_2, and 15 in t_3.

4.2 Connectivity matrices

The structural data collected is used in this thesis to build the connectivity matrices that are employed as input of CPM.

4.2.1 Preprocessing

All the preprocessing was done by using the Python library `ElikoPy` [58] (see Appendix A for more information). `ElikoPy` simplifies the processing of diffusion imaging by providing a standardized framework. The following corrections and steps were applied:

- *Brain extraction*: All the voxels corresponding to tissues surrounding the brain are suppressed in order to keep only the brain. It increases the speed and the efficiency of the following preprocessing steps [58].
- *Marchenko-Pastur Principal Component Analysis (MPPCA) Denoising*: Reduction of Rician noise. Using principal component analysis of the Marchenko-Pastur noise distribution gives a precise and efficient approach to evaluate and reduce noise. This algorithm does not blur the image nor create artifact [58].

- *Susceptibility field estimation*: Estimation of the susceptibility distortions [58].
- *Eddy and motion correction*: Correction of different artifacts (motion, susceptibility and Eddy current induced distortions) that affect the displacement and deformation of the brain [58].
- *White matter mask*: Applying a white matter mask accelerates microstructural features computation and helps doing tractography [58].

The results of the preprocessing steps can be seen on Fig.4.2 that represents a MRI scan before (left) and after (right) preprocessing.

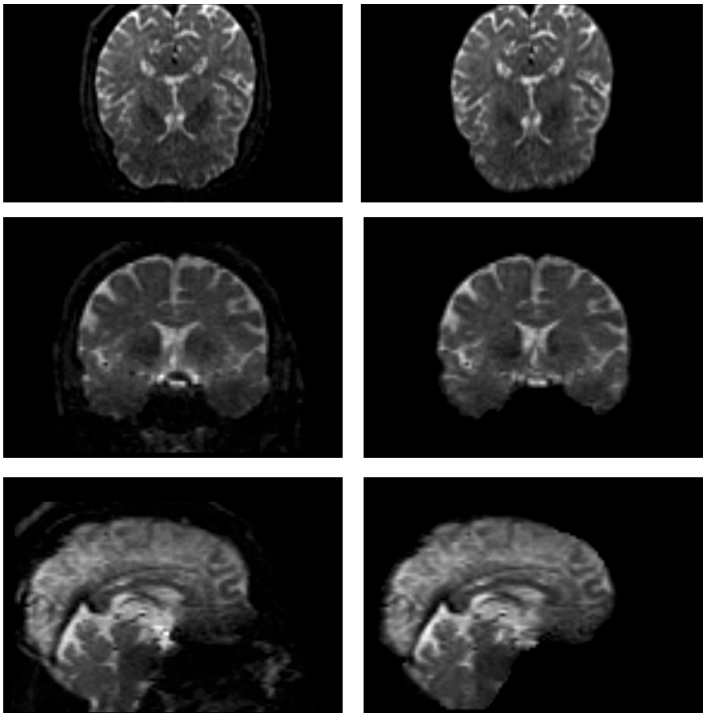


Figure 4.2: Comparison of a scan before (left) and after (right) preprocessing with `ElikoPy`.

DTI was also applied to register the scans and different maps in a common space. Moreover, multi-shell multi-tissue constrained spherical deconvolution (MSMT-CSD) is used to provide the tractograms. MSMT-CSD estimates the fiber orientation of diffusion-weighted MRI. It overcomes limitations of DTI, such as detecting multiple fiber orientations within a single voxel, as well as distinguishing different tissue types, and estimates a multi-tissue orientation distribution function (ODF) [59, 60].

4.2.2 Labeling

To know to which brain region each voxel belongs, parcellation and segmentation were done. Parcellation is the labeling of the cortical structures while segmentation is the labeling of the subcortical structures [61]. This was done using the `recon-all` command from the `FreeSurfer` package [62]. `FreeSurfer` is a widely used package designed to process, analyze, and visualize human brain MRI images. It is especially

used to extract anatomical information and produce brain surface models. The outputs of the `recon-all` command include among others cortical and subcortical volume measurements, cortical thickness estimates, representations of the boundary between the grey and white matters, and parcellation maps for several brain regions. While this command carries out many more processing steps than parcellation and segmentation, this thesis uses only those two outputs [63, 62, 64]. The *wmparc* map was employed since it possesses the highest number of labels of all the maps produced by `recon-all` (see Appendix B for the number of labels present in each map). It contains 177 different labels (see Fig. 4.3).



Figure 4.3: Wmparc map of a healthy participant. Each color corresponds to a different label, and thus to a different brain region.

4.2.3 Registration

Registration is a process in which a spatial transformation is found so that images in different spaces can be aligned in a common space (see Fig. 4.4) [65]. The labels maps, the different scans, and the FA maps obtained by applying DTI must be in the same space for each of the participants. If all those maps and scans were not properly aligned, the voxels in which the tracts pass through would be assigned to wrong labels. The labels maps were registered on the FA maps for each participant individually, and at each scan time.



Figure 4.4: Example of registration. Two scans that are in different spaces (left) are put in a common space (right). From [66].

The registration process contains two main steps: first finding an affine transformation that puts the moving file in the space of the static file, second applying the

mapping on the moving file. This was done using two functions, `getTransform` and `applyTransform` (inspired by [67]). The moving file was the *wmparc* maps and the static file the FA maps, to obtain registered labels maps in the anatomical spaces. By performing registration, potential issues related to scan size were effectively circumvented. The label maps were larger (256*256*256 voxels) than the anatomic scans and the FA maps. This inconsistency arose because `FreeSurfer` added additional voxels on each side to facilitate processing by using a cubic volume. However, registration introduces a *dilution phenomenon*. At the boundary between several brain regions, voxels of the new registered map overlap different regions of the original label map. An average is then applied to value and assign the voxels of the registered map to a brain region. This creates non-existent labels and distorts the boundaries between regions. An illustration can be seen on Fig. 4.5. To avoid this phenomenon, one label must be registered at a time. The mapping is applied on all the voxels of each label individually, then they are put back together.

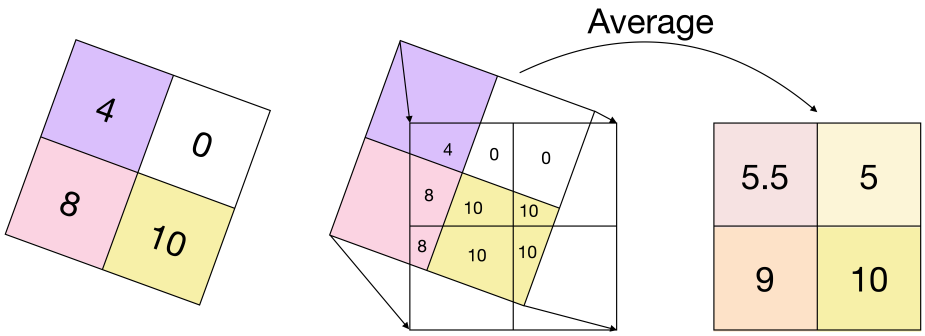


Figure 4.5: Illustration of the dilution phenomenon that occurs at the boundary between different brain regions. New non-existent labels appear at the right after the registration to a new space.

4.2.4 Structural connectivity

Tractography

As mentioned earlier, the construction of structural connectomes requires the use of tractography. The goal is to generate a complete tractogram of the entire brain. Probabilistic tractography was applied, which allows tracing fibers through regions characterized by high uncertainty. The `ElikoPy` function `tracking_solo` was used [67]. It provides tractograms based on the ODF map produced by the MSMT-CSD model, and allows to choose the number of streamlines of the tractogram, the maximum angle between two tractography steps, and the value below which streamlines do not propagate. Initial analyses were performed with the default values: 100000 streamlines, a maximum angle of 15°, and a cutoff of 0.1. However, after looking at the outcomes, it was observed that certain tracts ended up in the cerebrospinal fluid, indicating an abnormality. Consequently, the parameters were adapted. A maximum angle of 25° was taken, and the cutoff was kept at 0.1. The number of streamlines was increased to 250000 to have a more accurate representation of the brain connectivity. An example of tractograms obtained with these parameters for an AUD patient (top), and a control participant (bottom) are on Fig. 4.6.

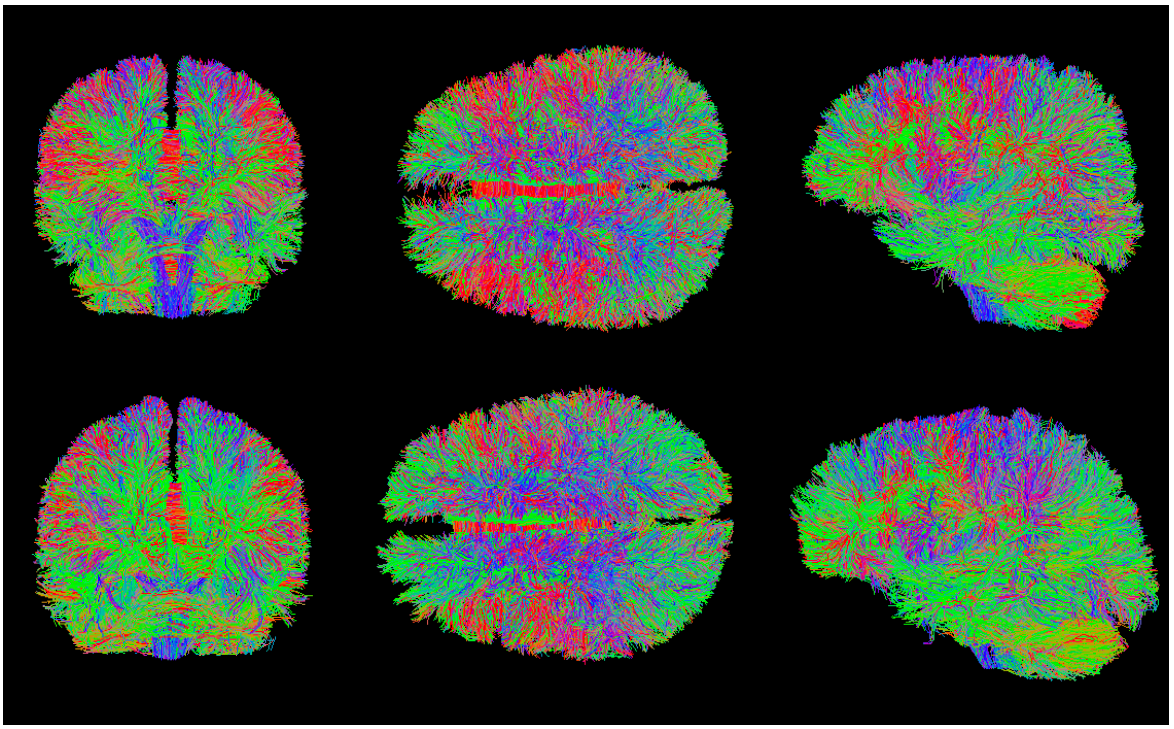


Figure 4.6: Tractograms for an AUD participant (top) and a control participant (bottom). Both tractograms contain 250000 streamlines, have a maximum angle between two tractography steps of 25°*, and a cutoff of 0.1.*

Connectivity matrices and normalization

The obtained tractograms served as input to build connectivity matrices. Each entry of the matrix corresponds to the number of fibers present between two brain regions. The connectivity matrices were computed using the function `utils.connectivity_matrix` of `DiPy` [68]. This function allows, if the `inclusive` argument is passed, to take into account streamlines passing through the different regions while constructing the matrix. Nevertheless, it requires a far greater computational time. To assess its potential impact on the outputs, it was applied to a small subset of participants. The observed changes were not significant, and given the additional computational time required, the argument was not used in this thesis.

An important step before building the connectivity matrices was confirming the presence of consistent regions across all patients. This verification was crucial, as the connectivity matrices need to be of uniform size for the CPM. Among the regions, the left and right vessels were not present in all patients. Since these two regions did not affect the objectives of this thesis, they were ignored for the patients in whom they were present. As mentioned earlier, defects were spotted in the tractograms, such as the presence of tracts ending up in cerebrospinal fluid. Tractography was therefore corrected. To make sure that such connections were absent from the connectivity matrices, the concerned regions were suppressed. The deleted regions are the cerebrospinal fluid, and the optic chiasm. Moreover, two more labels (right and left unsegmented white matter) were suppressed since they do not bring any additional information.

Initially, the construction of connectivity matrices involved considering the labels in ascending order. However, to enhance visual clarity and comprehensibility, the order in which the labels are taken into account was rearranged. The connectivity matrices are now organized as follows: labels from the left hemisphere, labels that are shared between both hemispheres and correspond to regions near the hemispheric junction, and finally labels from the right hemisphere. Within each hemisphere, the labels corresponding to grey matter were first taken into account, followed by the labels corresponding to white matter. Connectivity matrices before and after the sorting of the labels are illustrated in Fig. 4.7.

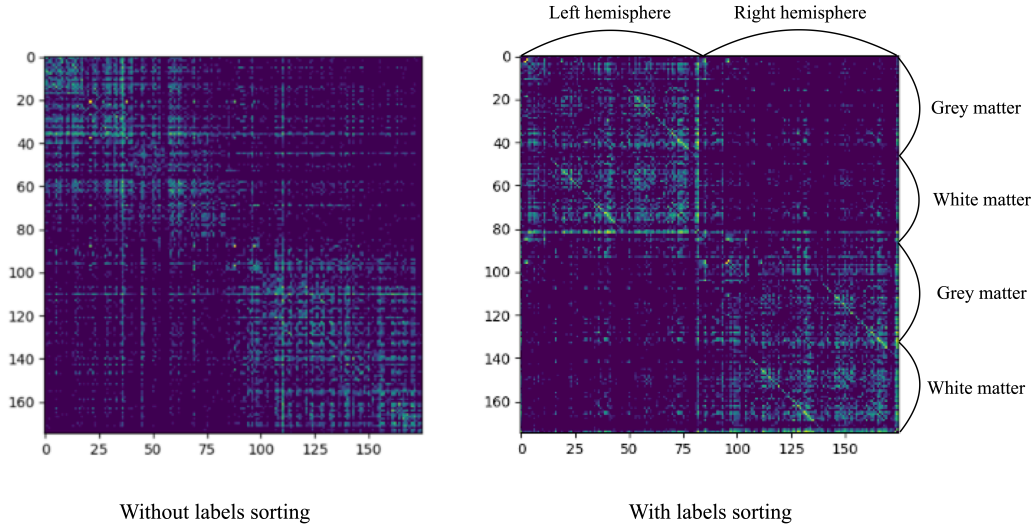


Figure 4.7: A connectivity matrix before (left) and after (right) the sorting of the labels.

However, connectivity matrices can not be used as is. Indeed, a normalization must be applied for several reasons. First, to compare matrices from different patients, the global size of the skull must be taken into account to avoid that the number of connections increases with the latter. Second, a bias is introduced by tractography. Indeed, the strength of connectivity decreases with the length of the tracts [69]. A first correction regarding the size of the different volumes was therefore applied. This is taken into account by dividing each entry of the matrix by the sum of the volumes of the two regions concerned [70, 71].

$$M_{ij} = \frac{M_{ij}}{\text{Volume}_i + \text{Volume}_j}, \tag{4.1}$$

with M the connectivity matrix, i the index of the first region concerned, and j the index of the second region concerned.

To correct the effects of fibers lengths, each entry of the connectivity matrix was multiplied by the ratio between the average length of the tracts between the two regions concerned and the average length of all fibers [72].

$$M_{ij} = M_{ij} * \frac{l_{ij}}{\bar{l}}, \tag{4.2}$$

with l_{ij} the average length of the fibers between the region i and the region j, and $\bar{l}$ the average length of all the fibers.

The effects of the different corrections applied can be seen on Fig. 4.8.

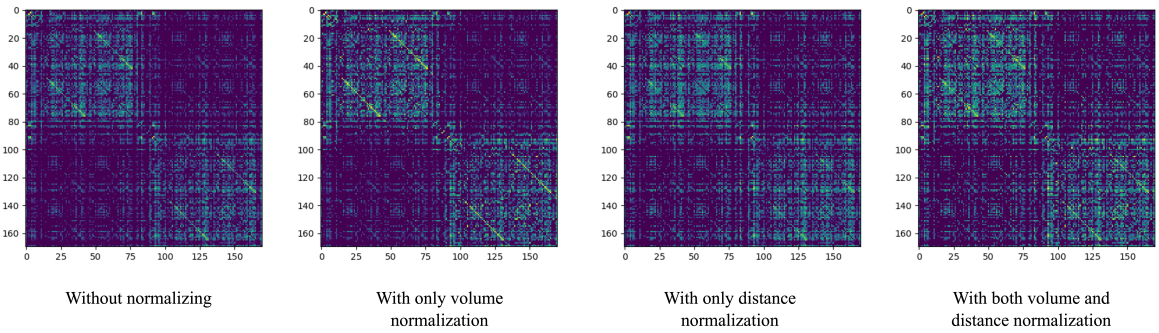


Figure 4.8: A connectivity matrix without any normalization (first image), with only the volume correction of eq. 4.1 (second image), with only the distance correction of eq. 4.2 (third image) and with both the volume and the distance correction (last image).

4.3 Connectome-based Predictive Model

The goal of CPM is to predict behavioral outcomes from connectivity data, using cross-validation. CPM takes into account the connectivity matrices of each participant, along with their corresponding behavior metrics. The prediction is performed individually for each behavior, following a series of steps:

1. Feature selection: Identification of relevant features from the connectivity data to be used for prediction.
2. Feature summarization: The selected features are summarized to obtain a representative measure for each participant.
3. Model building: A predictive model is built using the connectivity features and the corresponding behavior metrics.
4. Assessment of prediction significance: The significance of the predictions is evaluated to determine their reliability and statistical significance.

The different steps of CPM are illustrated in Fig. 4.9.

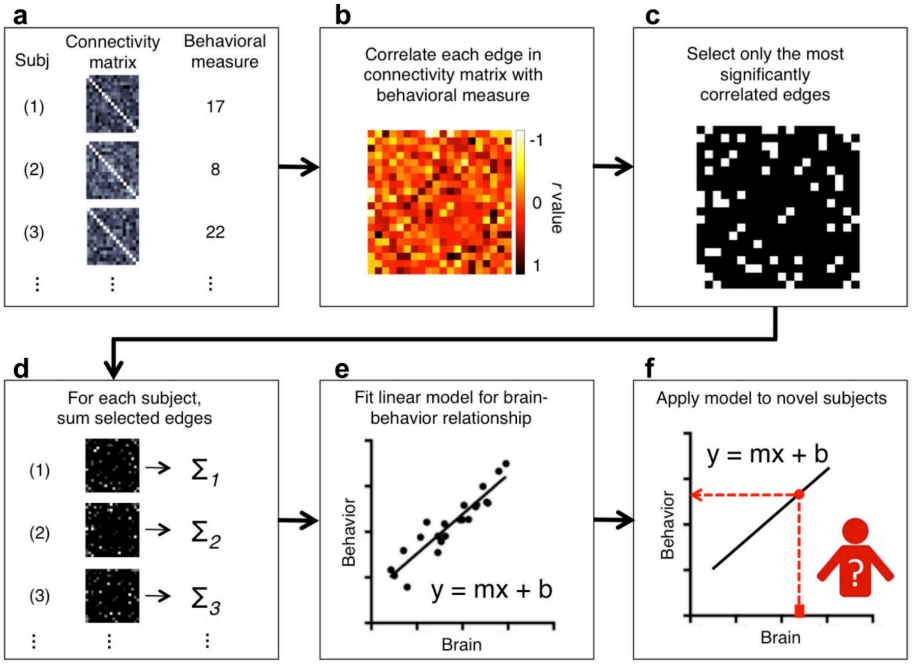


Figure 4.9: Schematic of connectome-based predictive modeling. a) The inputs of CPM are connectivity matrices and behavioral measures for each participant. It must be divided into a training set and a testing set. b) Correlation between each edge of the connectivity matrices across all participants in the training set and the behavioral measures. c) Selection of the most significantly correlated edges by using a threshold. d) Summarization of all the selected edges into a single value for each participant. It is done by summing all the edge strengths. e) Build a predictive model. f) Apply the predictive model on the testing set. From [3].

CPM was chosen because its approach with cross-validation allows to avoid overfitting and the lack of generalization present in more used methods such as correlation or regression models. Other methods such as multivariate-prediction or univariate regression methods also exist, but these are more computationally complex, and therefore were not chosen. Another alternative is Support Vector Regression, but this method bases its predictions on all present edges by giving them a specific weight. A method that selects more precisely the edges used for prediction was preferred, to extract biomarkers more easily [3].

4.3.1 Preliminary steps

Outliers

Outliers are data that deviate from the norm. It is important to detect and process them with specific care. Indeed, models trained on data with outliers can have reduced accuracy, increased error rates, or poor generalization. Identifying and treating outliers can thus improve model performance and lead to more reliable predictions [73].

At the behavioral level, we have defined outliers as participants having extreme variations in behavioral measures across the different measurements times. To identify

the outliers, boxplots were employed and can be found in Appendix C. The evolution of the various behavioral metrics is depicted in Fig. 4.10, where the red lines indicate the detected outliers. No outlier was detected for the compulsion metric. The participants detected as outliers were not included into the analysis.

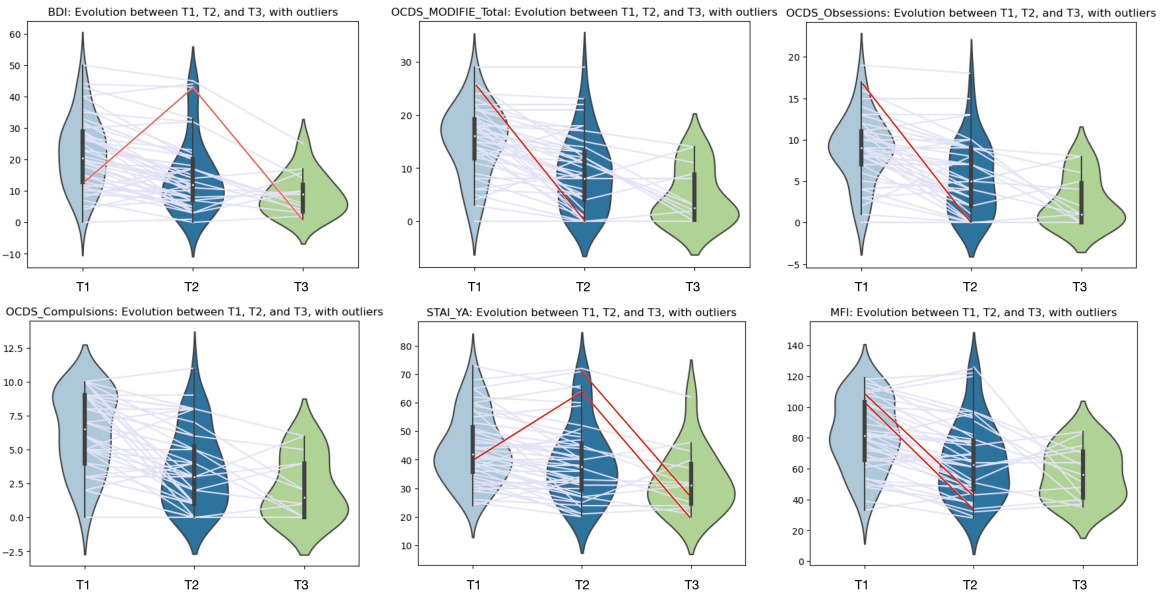


Figure 4.10: Evolution of the different behavior metrics for the AUD participants. The red lines are the participants that were identified as outliers.

Motion and behavior

The motion of the participant's head during data acquisition is known to be a potential source of confounding effects in structural data. Considerable head motion can generate false connectivity patterns. If there is a correlation between motion and the behavioral data, these motion-related patterns could increase prediction performance. That is why it is important to verify beforehand if motion and behavior metrics correlate, even if motion correction has already been applied while preprocessing the data. If it is the case, additional measures must be implemented, such as excluding high motion participants. In general, such participants are not considered suitable candidates in predictive models because the motion effects tend to persist even after motion correction [74].

Two motion metrics were explored: average absolute motion, and average relative motion. They were both produced by `ElikoPy` when doing the eddy and motion correction, and are measured in *mm*. The correlations between the different behavior metrics and both motion metrics were computed. The results can be found in Table 4.1.

Table 4.1: Correlations and p-values between the different behavior metrics, the average absolute motion, and the average relative motion of the participants. The p-value highlighted in red is the smallest p-value observed, and is close to the significance threshold.

	Absolute motion		Relative motion	
Behavior	Correlation	*p*-value	Correlation	*p*-value
BDI	0.075	0.364	0.158	0.056
OCDS	0.025	0.757	0.093	0.259
Obsessions	0.021	0.798	0.071	0.393
Compulsions	0.034	0.679	0.122	0.142
STAI	0.031	0.714	0.124	0.133
MFI	0.041	0.623	0.101	0.225

None of the *p*-values are considered significant since they all are bigger than 0.05. However, it is worth noting that the *p*-value between BDI and relative motion approaches the significance threshold of 0.05. This is nonetheless not the case for the absolute motion. This suggests that there is no significant correlation between motion and behavior, which is a favorable outcome. As a consequence, no correction was applied.

4.3.2 Model and parameters used

As stated earlier, CPM consists of four main steps: 1) feature selection, 2) feature summary, 3) model building, and 4) assessment of prediction significance. This section will go over the method and explain it more in details. Fig. 4.9 represents the key steps of CPM. Several combinations of parameters were explored, in order to identify the most effective set.

Load the inputs

CPM requires two inputs: connectivity matrices and behavioral data. For each patient, a connectivity matrix is generated for each scanning time, which corresponds to a specific behavioral data point. It is important to ensure the availability of both the scans and behavioral data for each patient and scan time, as data is missing. Only complete data for a given scan time and participant were processed. Due to variations in missing behavioral values and outliers across different behavioral metrics, the list of patients considered varies for each behavior.

Moreover, individual connectivity matrices are flattened and combined to form a comprehensive matrix containing all connectivity data. In this matrix, each row represents a specific edge across all participants, and each column represents the connectivity data for a single participant. The final matrix size is *(number of labels * number of labels) * number of participants*, where the number of labels equals 170 (after excluding the left and right vessels, the cerebrospinal fluid, the optic chiasm, the left and right unsegmented white matter, as well as the "unknown" label). This is represented in Fig. 4.9 as step *a*.

Division into training and testing sets for leave-one-out cross-validation

Cross-validation is a commonly used technique in machine learning and statistical modeling. It allows to assess the performance and the generalizability of a predictive model. It is designed to protect against overfitting, to test predictive model on new data, and helps guard against potential false positives. It involves partitioning the available data into multiple subsets or "folds" and iteratively evaluating the model's performance using different combinations of training and testing data. The leave-one-out cross-validation method was selected, in which only one measure is contained in the testing set. The model is then trained on the training set (containing all the other values), and its performance is evaluated based on how well it predicts the outcomes or targets of the testing data. The division into training and testing sets, as well as the training and evaluation of the model, is repeated, ensuring that each data point is used for testing and training. Cross-validation helps thus to address concerns of overfitting, which happen when a model performs well on the training data but fails to generalize to new, unseen data. By evaluating the model on multiple different subsets of data, cross-validation provides a more robust estimate of the model's performance and its ability to generalize to new data [75, 76, 77]. Leave-one-out cross-validation was preferred over other methods such as K-fold cross-validation to avoid doing the critical choice of the number of folds to perform. The choice of K impacts directly the variance and accuracy of the results obtained [78].

Relate connectivity to behavior

Each edge of the connectivity matrices is correlated to the behavioral measure. Two different correlation measures were tested: the Pearson correlation coefficient (r), and the Spearman rank correlation coefficient (r_S). The Spearman correlation coefficient is useful if there is a non-linear relationship between the connectivity data and the behavior measures or if the behavioral data do not follow a normal distribution, which may be the case. This step can be seen on Fig. 4.9 at the step *b*.

Feature selection

The most significant edges are selected based on the correlation coefficient computed at the previous step. The selection is done with a threshold, all the edges with a lower p-value than this threshold are selected. The threshold has thus an important impact on the results of the predictive model. It is therefore interesting to explore the effect of the threshold on the model accuracy. Different thresholds were experimented: 0.001, 0.01, 0.05, and 0.1.

Two different feature selections take place, a selection for the edges with a positive association with the behavioral measures, and one for the negative associations. Those two sets may be interpreted differently, and are obtained with the following equations:

$$\text{positive mask} = \text{edges with } \mathit{correlation} > 0 \text{ and } p - \mathit{value} < \mathit{threshold}, \tag{4.3}$$

$$\text{negative mask} = \text{edges with } \mathit{correlation} < 0 \text{ and } p - \mathit{value} < \mathit{threshold}. \tag{4.4}$$

This step is represented in the step *c* of the Fig. 4.9.

Feature summarization

For each participant in the training set, the selected edges are summarized to a single value per participant for the positive edge set and the negative edge set separately. This is done by summing all the edge strengths and dividing it by two since connectivity matrices are symmetric. The following equations are applied for each participant in the training set:

$$\text{positive summarization} = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\text{connectivity matrix}_{ij} * \text{positive mask}_{ij}, \tag{4.5}$$

$$\text{negative summarization} = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\text{connectivity matrix}_{ij} * \text{negative mask}_{ij}, \tag{4.6}$$

where n is the number of labels present in the connectivity matrices, and the positive masks and negative masks were obtained with the equations 4.3 and 4.4. This can be seen on step d of Fig. 4.9.

Model building and fitting

The model fitting is applied separately to both the positive edge set and the negative edge set. The standard CPM model assumes a linear relationship between the single-participant summary value and the behavioral variable. CPM may thus not be optimal for capturing more complex or non-linear associations between connectivity and behavior. To address this limitation, it is possible to enhance the model by incorporating higher-order polynomial terms. This may enhance the model's overall performance. However, using higher order polynomials may introduce overfitting. This can be seen on step e of Fig. 4.9.

In total, three different models were built. The first one, called in this thesis the *"Separate"* model, takes individually into account the positive edge set and the negative edge set. Both a linear regression and a second order polynomial fits were tried for this model. The linear regression is of the following form:

$$y_{linear} = \alpha * x + \beta, \tag{4.7}$$

where the coefficients α, and β are to be determined by the model fitting. The coefficients γ, δ, and ϵ for the following second order polynomial must also be determined by the model fitting:

$$y_{quadratic} = \gamma * x^2 + \delta * x + \epsilon. \tag{4.8}$$

Both equations 4.7 and 4.8 are applied individually on the positive edge set, and then on the negative edge set.

The second model, referred in this thesis as the *"Joint"* model, selects the positive and negative edges separately, but takes both into account when predicting the behavioral measures. Only a linear regression was used in this case since there are already two parameters in the model: the positive edge set and the negative edge set. The model is of the following form:

$$y_{joint} = \zeta * x_{pos} + \eta * x_{neg} + \theta.$$

Finally, the third model is referred as the *"Together"* model. When selecting the features, no distinction was done between the negative and the positive associations. Both a linear regression and a second order polynomial fits were also tried for this model. It thus follows the same form as equations 4.7 and 4.8 but is applied once on the positive and negative associations together.

The `NumPy` function `polyfit` was applied to fit the "Separate" and the "Together" model, and find the needed coefficients [79]. To do so, it minimizes the squared error:

$$\text{coefficients} = \min_{\text{coefficients}} \sum_{i=1}^{n} (train_behavior_i - y_i)^2,$$

where n is the number of participants in the training set, $train_behavior_i$ the behavioral measure for the ith participant of the training set, and y_i is the chosen model (linear or quadratic) applied on the summarization values of participant i, obtained with the equations 4.5 and 4.6. If the "Together" model is chosen, there is only one summarization value, accounting for the positive and negative associations. The `NumPy` function `lstsq` was used for fitting the "Joint" model, and follows the same logic while minimizing the Euclidean 2-norm [80].

Prediction on the participants of the testing set

After fitting the model on the training data set, computation of single-participant summary values for each participant in the testing set is done using the same method. For each participant in the testing set, this value was put into the predictive model estimated previously. This results in a predicted behavioral measure for the current participant. It can be seen on step f of Fig. 4.9.

Evaluation of the predictive model

A comparison between the predicted values and the true observed behavioral values is done to evaluate the accuracy of the predictive model. To do so, two measures can be used: correlation and mean squared error (MSE):

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2,$$

where n is the number of participants, Y_i the true behavioral measure of the participant i, and $\hat{Y}_i$ the predicted behavioral measure of the participant i. Correlation and MSE are usually linked, a higher correlation value indicates lower MSE and vice versa. Correlation provides an estimation of the relative predictive power, indicating how well the model can predict a new participant behavioral value within a known distribution. On the other hand, MSE provides an estimation of the absolute predictive power by quantifying the extent to which predicted values deviate from the true values. Correlation is commonly employed for functional MRI applications as a means to evaluate predictions of continuous variables obtained through cross-validation. Due to its scale ranging from -1 to 1, correlation coefficients offer an initial understanding of the model's performance. MSE lacks a specific scaling range and may not provide meaningful insights at first glance but is more used when the absolute accuracy of

predictions is needed. In this thesis, the primary focus is not on absolute prediction accuracy, which is why correlation was chosen as the evaluation metric. The final result of CPM is thus a correlation between predicted behavior and true behavior. These correlations obtained by using CPM are usually expected to lie between 0.2 and 0.5, as reported in [3]. Besides correlation, p-values are also computed to assess prediction significance. Only the p-values associated with a positive correlation were considered. Indeed, negative correlations mean that the predicted behavioral measures are opposite to the observed ones. These results are thus not usable.

Non-parametric assessment of prediction significance

Permutation testing is a complementary way to assess prediction significance, without strong assumptions on the underlying distributions. In this thesis, it involved maintaining the structure of the connectivity matrices while randomly reassigning the behavioral scores. The behavioral measures are shuffled, then CPM is applied, and the correlation coefficient is computed. This is usually repeated for 100 to 10000 iterations. In this thesis, 100 iterations were done since this step is computationally demanding. In the ideal case, all possible rearrangements of the data are performed, but this is not often done for computational reasons. However, it is important to ensure that an adequate number of iterations is employed to accurately estimate the distribution of the test statistic. Larger datasets may require a higher number of iterations to obtain reliable p-value estimates. The p-value of the permutation test is computed as the proportion of sampled permutations that are smaller to the prediction correlation without performing the permutations:

$$p\text{-}value = \frac{\text{Number of permutations with a correlation } < \text{ true predicted correlation}}{\text{Number of permutations performed}}.$$

The smaller the p-value, the better. In fact, this means that the p-value of the true results is one of the lowest compared to those obtained from the different permutations. It is therefore not reproducible with a random permutation.

4.3.3 Motion prediction

As mentioned earlier, motion impacts connectivity patterns. The relationship between motion and behavioral measures was already verified. However, it is also important to check if there is an association between motion and connectivity data. This can be done by looking whether CPM can predict motion or not. If it is the case, excluding high motion participants may be a solution. Partial correlation can also be used since it removes the effects of confounding factors.

Table 4.2: Results of the application of CPM to predict motion

	Predicted motion measures compared to the true ones	
Set of edges	Correlation	p-value
Positive	0.361	$8.987*10^{-6}$
Negative	0.500	$1.944*10^{-10}$

As it can be seen in Table 4.2, motion is predicted by CPM since the p-values are statistically significant. Partial correlation was thus used in the model to remove the effects of motion. The impact of using partial correlations has been evaluated. Results can be found in Appendix G

It is important to highlight that out of the 146 scans obtained, 67 of them are classified as having high motion, indicated by an average absolute motion greater than 0.15 mm [3]. Typically, participants with high motion are not considered ideal candidates for predictive modeling due to the potential persistence of nonlinear motion effects even after extensive motion correction. Therefore, it is generally recommended to exclude these participants from the dataset. However, in this particular thesis, it was decided to retain these participants to preserve a larger dataset, despite the presence of high motion in many cases.

4.3.4 Correcting for false positives

To ensure the absence of false positives, the *Bonferroni correction* is applied to the significance threshold for the p-values of the correlation evaluating the performance of the predictive model. The Bonferroni correction is a statistical technique that adjusts the p-value threshold for each individual test by dividing the desired significance level (α) by the total number of tests done. The usual value of the significance threshold is α=0.05. By implementing this adjustment, the Bonferroni correction reduces the overall probability of Type I errors (false positives) across all tests. However, it is a conservative method that may increase the risk of Type II errors (false negatives) [81]. In this thesis, a total of 40 different tests were conducted when exploring the different parameters used in CPM, and the significance level with the Bonferroni correction was recalibrated accordingly. The parameters used for each of the different tests can be seen in Appendix D. The new significance threshold becomes:

$$\alpha_{Bonferroni} = \frac{\alpha}{\text{Number of tests conducted}} = \frac{0.05}{40} = 0.00125.$$

Another method of correction of the significance threshold can also be used to overcome the increasing risk of false negatives that comes with the Bonferroni method. This second method is the *Benjamini-Hochberg procedure*, which also ensures the absence of false positives by controlling the false discovery rate. It consists of ranging all the p-values obtained in increasing order, and giving to each p-value a number i corresponding to its place in the ranked list. The smallest p-value has a value of $i = 1$ while the biggest p-value has a value $i = m$, m being the total number of tests performed. Each p-value is then compared to the Benjamini-Hochberg critical value, stated in the following equation:

$$\alpha_{Benjamini-Hochberg} = \frac{i}{m}Q,$$

where i is the rank of the concerned p-value, m the total number of tests performed, and Q is the false discovery rate previously selected. The biggest p-value being strictly inferior to $\alpha_{Benjamini-Hochberg}$ sets the threshold. All the p-values smaller than the threshold are considered significant, even if they are not smaller than their Benjamini-Hochberg critical value. The choice of Q is critical, since it sets the value of the highest

p-value that may be considered as significant. If the cost of conducting additional experiments is relatively low and the risk of overlooking a potentially significant finding is high, it is advisable to employ a relatively high false discovery rate, such as 0.1. This ensures that important discoveries are not missed, prioritizing sensitivity over specificity [82]. Q was set to 0.05 in this thesis since many p-values obtained are below 0.05.

4.4 Final models selection

40 variations of CPM ran for each behavior. Indeed, different sets of parameters were tried to spot the most suitable ones. The parameters that varied are:

- Correlation used for the feature selection: Pearson or Spearman rank correlation,
- Threshold used for the feature selection: 0.001, 0.01, 0.05, or 0.1,
- Model: "Separate", "Joint", or "Together",
- Degree of the polynomial fit: linear or quadratic.

Three different sets of parameters were selected for each behavior. First, the parameter set with the most favorable outcomes solely for positive associations was picked ("Separate" model for the positive associations). The same was done for the parameter set with the best results exclusively for negative associations ("Separate" model for the negative associations). Finally, the optimal parameter set for the best of the two models considering both positive and negative associations ("Joint" and "Together" models) was determined.

4.5 Extracting the biomarkers

The last step is identifying biomarkers using the results obtained from CPM for the selected models. It is likely that a slightly different set of features is selected in each iteration of cross-validation. However, the correct features should overlap across all iterations. The most cautious approach is to consider only the edges that were consistently selected in all iterations, indicating their reliability. Alternatively, a less stringent threshold can be set, allowing edges to be included if they appear in most iterations, providing a slightly more flexible criterion. The first option was chosen.

Since many significant edges were identified, an additional step was performed to refine the selection process. For each selected model and behavior, only the 20 edges with the strongest average correlation with the behavioral measure were picked for further analysis.

The five most frequently occurring regions (nodes) for each behavior were also selected by analyzing the number of times each region appeared in the significant connections.

Chapter 5

Results

5.1 Behavioral changes during abstinence

The different behavioral measures evolve for the participants suffering from AUD during abstinence. This evolution can be seen on Fig. 5.1. The red boxes on the figure represent the range in which the behavioral measures of control participants lie. Control individuals experience minimal (scores between 0 and 9) to mild depression (scores between 10 and 18), while the majority of the participants with AUD suffers from moderate depression (scores between 19 and 29) at the time t_1. A few subjects experience severe depression (scores above 30). The Obsessive Compulsive Drinking Scale (OCDS) assesses a global craving score, an obsession score, and a compulsion score. The highest the scores, the more cravings, obsessions, and compulsions are experienced. At the time t_1, the AUD participants have higher scores than controls across the three categories. For state and trait anxiety (STAI), the values for the control participants are more spread. At is also the case for the values of the AUD participants. A high score indicates a high level of anxiety. Finally, the majority of the AUD participants have bigger multidimensional fatigue scores (MFI) than control subjects.

During the period of abstinence, a general decline in the various behavioral measures is observed. Interpreting changes between t_2 and t_3 is more challenging than doing the same for t_1 and t_2 due to the limited number of available measures. Additionally, there are noticeable variations in the distribution of scores among different participants across the scan times. Specifically, for the BDI measure, scores were widely dispersed at t_1, but the majority decreased at t_2, with only a few individuals maintaining high values. In contrast, the other behavioral measures remained scattered even after a three week period of abstinence.

When comparing the behavioral data of alcohol-dependent patients with that of control participants, it can be observed that the majority of data collected from patients suffering from AUD in t_3 for BDI, OCDS, Obsessions, and Compulsions lie within the same range as the control participants. However, it is challenging to draw comparisons between alcohol dependent participants and control participants for STAI and MFI measures since control participants exhibit wide variations of the values.

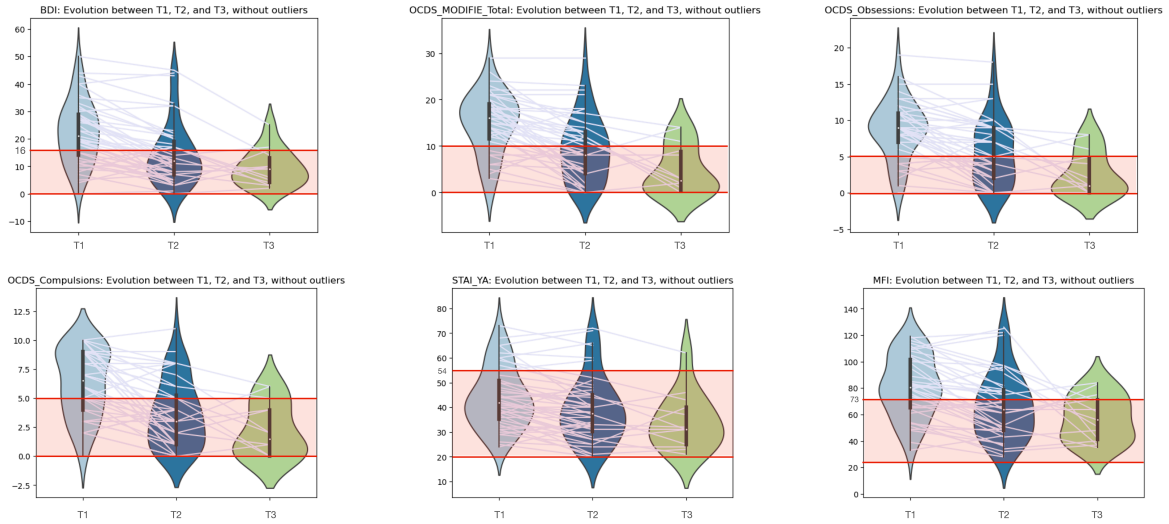


Figure 5.1: Evolution of the different behavioral measures at time t_1, t_2 and t_3, without the outliers that were identified. The red boxes represent the the range of the controls' behavioral measures.

Moreover, some individuals exhibit an increase in measures either between time points t_1 and t_2, or between t_2 and t_3. It is worth noting that 16 participants drank during the abstinence period, which may affect the evolution of the different behavioral measures.

5.2 Best fitting CPM models

5.2.1 Selected parameter sets

As mentioned in the methods, different set of parameters were explored for the distinct models. Those sets yielded to different predictions and results. This subsection reviews the best sets of parameters, globally and for each individual behavior.

Out of the 40 tested models, the best three are selected for the identification of the biomarkers. Those models are the ones with highest predictive value of the behavioral data, assessed by the p-value, separately for positive associations ("Separate" model), negative associations ("Separate" model), and both ("Joint" or "Together" models).

Overall, the parameter set with the smallest p-value across all behaviors for positive associations uses a Pearson correlation coefficient (r), a threshold of 0.1, and a linear fit. For negative associations, Pearson correlation coefficient was used with a threshold of 0.001 to select the edges, and a linear fit was employed. Finally, the "Joint" model was selected with Pearson correlation coefficient and a threshold of 0.1 for the selection of the edges have the smallest p-values across the behaviors. However, distinct sets of parameters are yielding to higher predictive power for specific behavioral measures. The parameters sets that have the best results for each category and each behavior individually are in Table 5.1. Those are the ones that will be used for further analyses.

Table 5.1: Selected sets of parameters with the smallest p-values for each behavior, and their prediction performance. Pearson correlation coefficient is represented by r, and Spearman correlation coefficient by r_S. The p-values highlighted in red are above the Bonferroni threshold.

		Parameters used				Results	
Behavior	Category	Model	Coeff.	Thresh.	Order	Corr.	p
BDI	Positive	Separate	r_S	0.01	1	0.27	0.002
	Negative	Separate	r	0.001	2	0.26	0.002
	Both	Joint	r	0.1	1	0.28	0.001
OCDS	Positive	Separate	r_S	0.001	1	0.29	0.001
	Negative	Separate	r	0.001	2	0.34	9e-05
	Both	Joint	r_S	0.001	1	0.37	1e-05
Obsessions	Positive	Separate	r	0.01	1	0.33	1e-04
	Negative	Separate	r	0.001	2	0.37	2e-05
	Both	Joint	r	0.001	1	0.36	2e-05
Compulsions	Positive	Separate	r_S	0.001	1	0.25	0.004
	Negative	Separate	r	0.001	1	0.40	2e-06
	Both	Joint	r_S	0.001	1	0.36	2e-05
STAI	Positive	Separate	r	0.1	1	0.25	0.003
	Negative	Separate	r_S	0.01	2	0.32	2e-04
	Both	Joint	r	0.1	1	0.39	5e-06
MFI	Positive	Separate	r_S	0.05	1	0.26	0.002
	Negative	Separate	r	0.001	2	0.37	1e-05
	Both	Together	r_S	0.001	2	0.36	2e-05

The correlations obtained with those sets of parameters range from 0.25 to 0.40 (see Table 5.1), which is in the expected range (0.2 to 0.5) stated by Shen et al. in their study [3]. Those are also the highest correlations observed across all tested models (see Appendix E). Some of the p-values obtained are above the threshold the Bonferroni corrected significance threshold. They are represented in red in the Table. As mentioned earlier, it is the case for both "Separate" models for BDI, as well as for the positive associations of the "Separate" model for Compulsions, STAI, and MFI. However, all those p-values are considered significant with the Benjamini-Hochberg procedure. The Obsessions score has particularly high correlations across the different models used while BDI has the smallest values considering all the different categories. The highest correlation obtained is for the "Separate" model applied to the negative associations of Compulsions.

5.2.2 Permutation testing

Permutation testing was performed for the selected models. The permutation-based p-values obtained for each behavior and model are shown in Table 5.2. The values highlighted in red are above the significance threshold of 0.05. All those correspond to models that were above the Bonferroni threshold (see Table 5.1). As a consequence and to ensure the significance of the results, those models were not taken into account for biomarkers.

For Obsessions, Compulsions, STAI, and MFI, some of the p-values of the permutation testing equal exactly zero. It is due to the small number of permutation performed, a 100 per model.

Table 5.2: P-values obtained with the permutation testing for each behavior and model selected. The values highlighted in red are above the significance threshold of 0.05.

Behavior	Pos. associations	Neg. associations	Both
BDI	0.009	0.188	0.019
OCDS	0.059	0.019	0.019
Obsessions	0	0.029	0.009
Compulsions	0.089	0	0
STAI	0	0	0
MFI	0	0.029	0.009

5.3 Biomarkers

As said in section 4.5, three different models with their own set of parameters were selected to identify the biomarkers for each behavior: the "Separate" model with the smallest p-value for positive associations, the "Separate" model with the smallest p-value for negative associations, and the combined ("Joint" or "Together") model with the smallest p-value. This led to different significant edges being identified for each behavior. Table 5.3 contains the number of selected edges for each of the models.

Table 5.3: Number of significant edges identified for each of the models with the best prediction accuracies, and for each of the behaviors. The values in italic are the cases where the exact same edges are identified by both models.

	Number of edges selected for different models			
	"Separate" model		Combined model	
Behavior	Positive edges	Negative edges	Positive edges	Negative edges
BDI	53	2	432	166
OCDS	*8*	6	*8*	20
Obsessions	120	*11*	20	*11*
Compulsions	*6*	4	*6*	20
STAI	*418*	51	*418*	209
MFI	413	6	13	20

For each of the behaviors, the significant edges have been compared across the different models. In all cases, there is a significant overlap of identified edges across the models. Even more, in some cases, the overlap is total, emphasizing the relevance of those edges for predicting behavior. For example, the same 8 positive edges were selected for OCDS by both the separate and combined models. Other cases are highlighted in Table 5.3.

The detected edges are distributed differently between the hemispheres of the brain, as well as between the types of structures they connect, as described in Tables 5.4 and 5.5.

For BDI, Obsessions, STAI, and MFI, connections within the same hemisphere are more frequent (see Table 5.4). For OCDS and Compulsions, there are more connections linking the two hemispheres. Comparing both hemispheres, only OCDS shows more connections within the right hemisphere than the left, while it is the contrary for all the other behavioral measures. Moreover, it is worth noting that the Obsessions behavior includes more connections involving the central parts of the brain such as the corpus callosum and the brain stem, labeled as "middle" in Table 5.4.

Table 5.4: Number of edges linking the different hemispheres of the brain for each behavior.

	BDI	OCDS	Obs.	Comp.	STAI	MFI
Left – Left	33.7%	9.1%	25.9%	17.9%	32.6%	38.4%
Right – Right	27%	21.2%	21.5%	14.3%	21%	25.2%
Left – Right	30.9%	51.5%	27.4%	53.6%	37.5%	25.9%
Middle – Middle	0.5%	0%	1.5%	0%	0.3%	0.5%
Middle – Left	3.9%	9.1%	11.9%	7.1%	4.3%	5%
Middle – Right	3.1%	0%	8.9%	3.6%	2.1%	4.1%

Regarding the types of structures linked, the connections are also distributed differently depending on the behavior assessed (see Table 5.5). For BDI, Obsessions, Compulsions, STAI, and MFI, most edges connect white matter to cortex. For OCDS, and Compulsions, the other structure types play an essential role, being involved respectively in 57.6% and 50% of the connections. Edges between white matter structures are important for BDI.

Table 5.5: Number of edges linking the different structures of the brain for each behavior.

	BDI	OCDS	Obs.	Comp.	STAI	MFI
White Matter – White Matter	22.1%	12.1%	17%	14.3%	16.6%	15.8%
Cortex – Cortex	11.3%	9.1%	5.9%	3.6%	12.8%	16.9%
White Matter – Cortex	31.4%	12.1%	23%	28.6%	35%	29.5%
Other – Other	7%	0%	20%	0%	5.4%	6.6%
Other – White Matter	15.2%	30.3%	16.3%	25%	14.2%	16.7%
Other – Cortex	12.1%	27.3%	14.8%	25%	13.8%	13.5%

5.3.1 Selected connections

Some edges play a role for different behaviors. The following edges are present for at least five of the six behaviors assessed:

- Left superior parietal cortex and the right thalamus proper: present for all,

- Left thalamus proper and right precuneus cortex: present for BDI, OCDS, Obsessions, STAI, and MFI,
- Right hippocampus and right entorhinal white matter: present for OCDS, Obsessions, Compulsions, STAI, and MFI,
- Left superior parietal cortex and right ventral diencephalon: present for BDI, OCDS, Obsessions, Compulsions, and STAI.

BDI

All the connections selected for BDI are in Table F.1 of Appendix F. The negative associations found with the "Separate" model are the same as the ones obtained with the "Joint" model. 62.3% of the selected edges are positively associated connections while the remaining 37.7% are negatively associated connections. 21.2% of the positively associated connections were selected by both the "Separate" and the "Joint" model.

OCDS

All the connections selected for OCDS are in Table F.2 of Appendix F. The positive associations found with the "Separate" model are the same as the ones obtained with the "Joint" model. Interestingly, only 26.7% of the selected edges correspond to positively associated connections, while the majority (73.3%) are negatively associated connections. 18.2% of the negatively associated connections were selected by both the "Separate" and the "Joint" model.

Obsessions

All the connections that were selected for Obsessions can be found in Table F.3 of Appendix F. The negative associations identified by the "Separate" model align with those obtained by the "Joint" model. 37.9% of the positive associations edges were selected in both the "Separate" and the "Joint" models. 72.5% of the selected edges correspond to positively associated connections, while 27.5% are negatively associated connections.

Compulsions

Table F.4 of Appendix F displays all the connections selected for Compulsions. The same positive associations were selected by the "Separate" model and the "Joint" model. Many negatively associated connections were selected: 77.8% of the total number of edges kept.

STAI

Table F.5 of Appendix F includes all the connections that were selected for STAI. The positive associations of both the "Separate" and "Joint" models have chosen the same connections. In terms of the distinction between positively associated and negatively associated connections, it is less prominent in this case. 62% of the selected

connections are indeed negatively associated, while the remaining 38% are positively associated.

MFI

The connections selected for MFI are in Table F.6 of Appendix F. All the negative associations edges chosen by the "Separate" model were also selected by the "Together" model. The opposite is true for positive edges, all the positive associations edges selected by the "Together" model were also chosen by the "Separate" model. There are as many negatively associated connections as positively associated ones.

Comparison between OCDS, Obsessions, and Compulsions

Since the OCDS test provides three related scores (total craving, obsessions and compulsions), it is therefore interesting to compare the edges selected for these three measures. First of all, three edges are shared by the three measures: the edges between the right hippocampus and the right entorhinal white matter, between the left superior parietal cortex and the right choroid plexus, as well as with the right ventral diencephalon. 15 connections are shared between OCDS and Compulsions, while only six are common to OCDS and Obsessions. Obsessions and Compulsions share none of their selected edges. Moreover, 31 edges are specific to Obsessions, six to OCDS, and nine to Compulsions.

5.3.2 Regions as biomarkers

The goal of this section is to identify which regions have the greatest impact on the prediction of each behavioral measure. Those are our biomarkers.

The five most recurring regions for BDI are the right superior parietal white matter, the left inferior parietal white matter, the left superior parietal white matter, the left lateral occipital white matter, and the left transverse temporal white matter. These can be seen on Fig. 5.2. The connections between the left lateral occipital white matter, and the right superior parietal white matter, the left inferior parietal white matter and the left superior parietal white matter were selected for BDI as significant. The two last regions are also connected to each other, and this connection is also significant in predicting BDI.

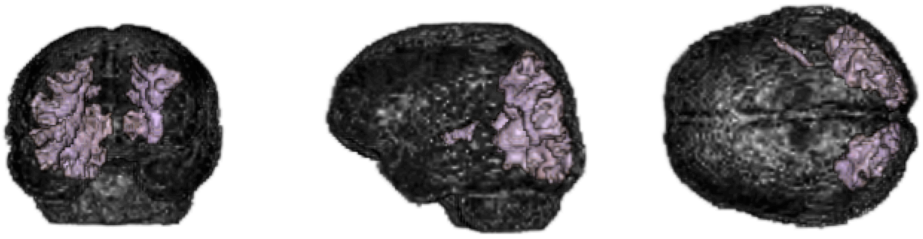


Figure 5.2: Anterior, lateral, and superior views of the regions selected for BDI.

Regarding OCDS, the selected regions are the left superior parietal cortex, the brain stem, the left precuneus cortex, the right choroid plexus, and the right banks of the superior temporal sulcus. These regions are showed on Fig. 5.3. The significant selected connections only contain an edge between these regions (between the right choroid plexus and the left superior parietal cortex).

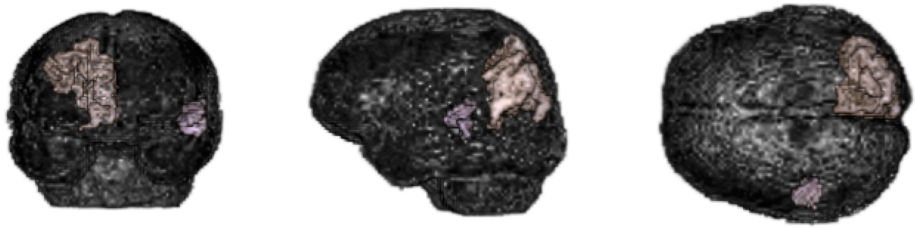


Figure 5.3: Anterior, lateral, and superior views of the regions selected for OCDS.

In regards of Obsessions, the brain stem is the most frequently occurring region, followed by the mid posterior corpus callosum, the left superior temporal cortex, the right insular cortex, and the left thalamus proper. These can be seen on Fig. 5.4. The connections between the brain stem and the left thalamus proper, the right insular cortex, and the mid posterior corpus callosum were selected as important for predicting Obsessions. Similarly, the connections between the left thalamus proper and the right insular cortex, as well as with the mid posterior corpus callosum, were also found to be significant. However, no edge was observed between the left superior temporal cortex and the other five most recurrent regions in the selected connections for Obsessions.

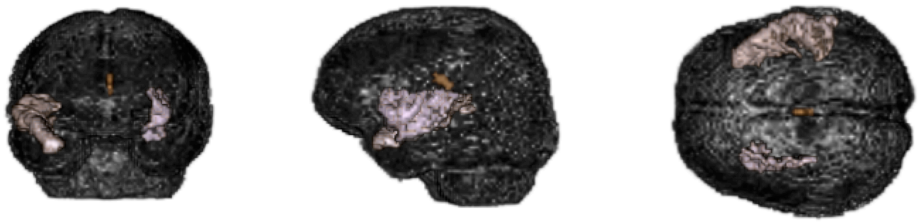


Figure 5.4: Anterior, lateral, and superior views of the regions selected for Obsessions.

For Compulsions, the five most recurring regions are the left superior parietal cortex, the left precuneus cortex, the brain stem, the right hippocampus, and the left transverse temporal white matter. The selected regions are showed on Fig. 5.5. However, in the selected edges for predicting Compulsions, none of these regions are found to be connected to each other.

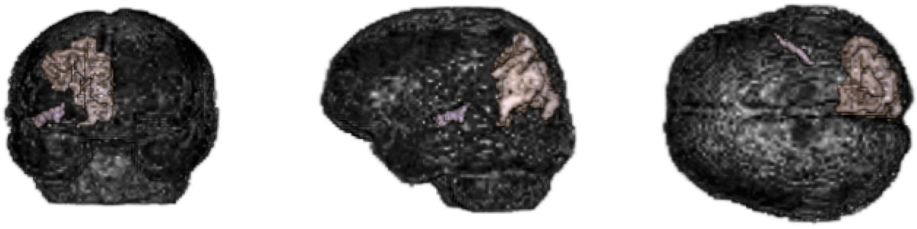


Figure 5.5: Anterior, lateral, and superior views of the regions selected for Compulsions.

The selected regions for STAI are the left superior temporal cortex, the left temporal pole white matter, the left lateral occipital white matter and cortex, and the left temporal pole cortex. These can be seen on Fig. 5.6. The connections between the two regions of the temporal pole and the lateral occipital regions were selected to predict STAI. It is also the case for the connections between the left superior temporal cortex and the lateral occipital regions.

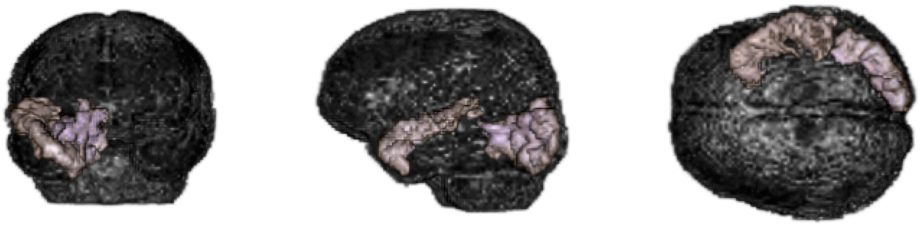


Figure 5.6: Anterior, lateral, and superior views of the regions selected for STAI.

Finally, for MFI, the regions are the left inferior parietal white matter, followed by the left paracentral white matter, the left superior temporal white matter, the left temporal superior cortex, the left superior parietal white matter, and the left temporal pole cortex. These regions are showed on Fig. 5.7. The selected edges for predicting MFI contains all the connections between these regions, except the connection between the left paracentral white matter and the left temporal pole cortex.

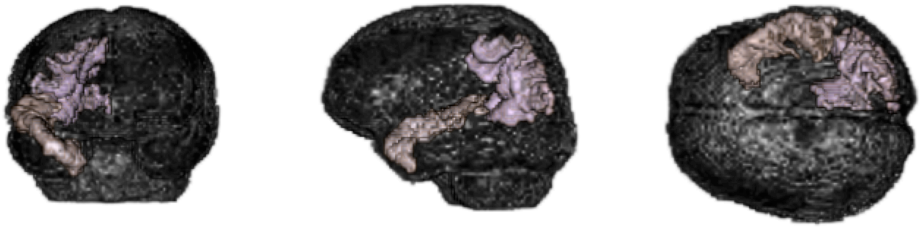


Figure 5.7: Anterior, lateral, and superior views of the regions selected for MFI.

5.4 Factors affecting CPM predictions

5.4.1 Effects of the different parameters

In this section, the impact of the sets of parameters on the predictive performance of the 40 explored models is reviewed.

The correlations and p-values obtained for the different models and parameters are in Appendix E. Overall, the correlations produced range from -0.58 to 0.40, and the p-values are between 1.6e-06 and 0.94. The wide ranges observed indicate that the accuracy of predictions varies across the different models.

Models

The "Together" model is the least predictive for the various behavioral measures, with correlations between -0.58 and 0.36. Indeed, only 14% of tests performed have a

p-value below the significance threshold of 0.05 (range: 2.08e-05 to 0.94), and merely 2% of tests are below the Bonferroni-corrected threshold.

In the "Separate" model, 75% of the predictions associated with positive associations have p-values below 0.05, while 14% are below the Bonferroni threshold. Similarly, for negative associations in the "Separate" model, 75% have p-values below 0.05, and 20% fall below the Bonferroni-corrected value.

Notably, the "Joint" model demonstrates the highest predictive power, with 93% of p-values being below 0.05 and 48% below the Bonferroni threshold. The correlations range from 0.15 to 0.39 and the p-values from 4.7e-06 to 0.09.

It is also worth noting that one behavioral measure was less well predicted than others across all models and parameter sets tested. In fact, the results were conclusive for all behavioral measures except for BDI, for which only 1.7% of the p-values are below the Bonferroni threshold, despite having 45% below 0.05. The behavior with the most p-values below the Bonferroni threshold is Obsessions. The percentage of results associated with a p-value below 0.05 and below the Bonferroni threshold are gathered for each behavior in Table 5.6.

Table 5.6: Percentage of p-values obtained across the different models that are below 0.05 and below the Bonferroni thresholds for the different behavioral measures.

	Percentage of the p-values below ...	
Behavior	0.05	the Bonferroni threshold
BDI	45%	1.7%
OCDS	65%	23.3%
Obsessions	71.7%	31.7%
Compulsions	50%	13.3%
STAI	48.2%	12.5%
MFI	56.7%	15%

Even though the "Separate" model predicts well globally for both the positive and negative associations, its performance varies drastically depending on the behavioral measure considered. It did not produce any p-values below the Bonferroni threshold for the positive associations related to BDI, Compulsions, STAI, and MFI.

Threshold for selecting the edges

Different thresholds were used for edges selection: 0.001, 0.01, 0.05, and 0.1. In general, the smallest threshold obtained the best results. Indeed, 51% of the p-values obtained while using it are below the significance threshold of 0.05, 27.5% are below the Bonferroni threshold, and 61.1% of the best parameters sets use this selection threshold. The results for all the other thresholds are similar to one another. 61.9% to 65.4% of the p-values obtained while using a threshold of 0.01, 0.05, or 0.1 are below 0.05, and 12% to 16% are below the Bonferroni threshold.

Correlation used

Two different correlation methods, Pearson correlation coefficient and Spearman rank correlation coefficient, were employed to examine the relationship between the connectivity matrix edges and behavioral measures. Both approaches yielded similar results. Specifically, 66.4% of the p-values obtained through Pearson correlation were below 0.05, while 52.9% were below 0.05 when using Spearman rank correlation. Moreover, 20.1% of the p-values fell below the Bonferroni threshold for Pearson correlation, and 14.2% for Spearman rank correlation.

Degree for the model building and fitting

The models used can be classified into three categories based on their degrees: linear models (degree 1), quadratic models (degree 2), and the "Joint" model (also degree 1 but considering both positive and negative associations). Among these categories, the "Joint" model stands out in terms of the number of significant p-values obtained. Notably, 93.7% of the p-values obtained using the "Joint" model are below 0.05, and 47.9% are below the Bonferroni threshold. In comparison, for the linear and quadratic models, 53.2% and 56.3% of the p-values are below 0.05, and 13.3% and 11.2% are below the Bonferroni threshold, respectively.

Additionally, it is worth noting that for each type of model employed ("Separate" and "Together"), the choice between a linear fit or a quadratic fit consistently led to comparable outcomes. Specifically, the p-values for the different behavioral measures were below the same threshold for the same parameters, regardless of whether a linear or quadratic degree was used.

Chapter 6

Discussion

In this thesis, we applied CPM to extract biomarkers of behavioral changes during abstinence. It is, to our knowledge, one of the first times it is used on structural data. Prediction results, measured by the correlation coefficients between the true behavioral values and the predicted ones range from 0.252 to 0.404. This is consistent with Shen et al., who stated that correlations obtained with this method are expected to lie between 0.2 and 0.5 [3]. This suggests that CPM can effectively be used on structural MRI data.

Using this method, the biomarkers selected show the global prevalence of the left hemisphere to predict the behavioral measures. As expected, they also highlight that each behavior is linked with different regions. Those are summarized in Fig. 6.1.

6.1 Biomarkers

6.1.1 Connections

Some connections that were identified in the state of the art as being impacted by AUD were also selected by the different models. For the *Executive Control Network*, connections between parietal regions and temporal regions were picked for BDI, Obsessions, STAI, and MFI. The findings in the literature stated that such connections were negatively correlated with failed alcohol consumption control. The OCDS measure is the one that comes closest to the concept of alcohol consumption control. However, no link between failed alcohol consumption control and these edges was found in this thesis, as those connections were not selected for OCDS. BDI, Obsessions, STAI, and MFI also presented connections belonging to the *Attention network*. In fact, some selected connections link superior parietal regions to temporal regions. STAI and MFI exhibit connections belonging to the *Somatosensory network*, between postcentral, temporal and parietal regions. The *corpus callosum* was identified for all the behaviors. Finally, MFI presents a connection between the *cerebellum* and the frontal cortex.

Moreover, the number of edges selected by the different models seems to indicate that some behaviors are more difficult to predict. BDI and STAI are the behavioral measures with the highest number of edges selected, and with the highest p-values regarding the positive and negative associations. This may suggest that a potential overfit, and that more edges are needed to obtain significant results.

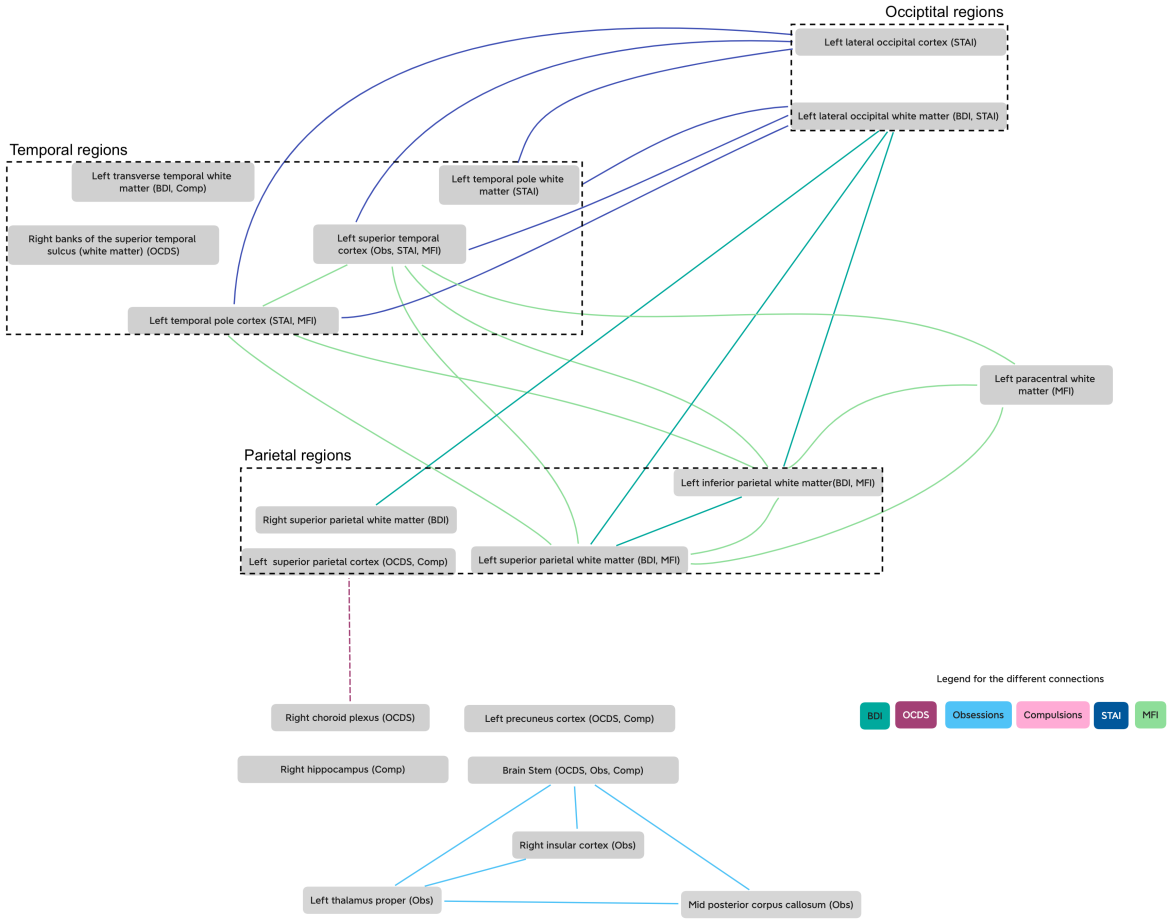


Figure 6.1: Illustration of the different regions selected for each of the behaviors, and their connections to one another. The color of the connection indicates for which behavioral measure the connection is significant. The plain lines signify that the connection is positive, while the dashed lines are negative. The behaviors indicated in the brackets for each region indicate for which behavior the region was selected.

6.1.2 Regions

This section reviews the different regions selected, their functions, the links with the literature review done in the Part I, as well as their links with the different behaviors concerned found in the literature.

The majority of these regions are mentioned in the Part I as being impacted by AUD. Surprisingly, the left lateral occipital cortex, the left lateral occipital white matter, and the left paracentral white matter were not identified in the literature as having structural connectivity impairments due to AUD.

All the regions selected for each of the behaviors, as well as their connections to one another are represented on Fig. 6.1. It is worth noting that none of the regions selected for Compulsions are connected to one another, and that only one connection is present for OCDS. Moreover, this connection is the only negative one. Interestingly, the chosen regions also have negative connections with other regions of the brain, but these were not picked as the most recurring regions. Since the "Joint" or "Together" models selected many negative edges, it suggests that the positive edges have a stronger predictive impact than the negative edges.

Temporal regions

Different regions belonging in the temporal region were found.

First, the left superior temporal cortex was selected as a positively associated region for Obsessions, STAI, and MFI. Given that the *superior temporal region* is known to be altered in teenagers with AUD, it is logical that the left superior temporal cortex has an impact on different behaviors. The superior temporal region has various functions, such as integration of different sensory inputs, language processing, speech comprehension, spatial processing and higher-level cognitive functions. It plays a role in memory, attention, social information analysis, and perception of emotions in facial stimuli [83, 84, 85, 86, 87].

Studies showed that the superior temporal gyrus may be involved in OCDS, and thus may impact Obsessions. It plays a role in visuospatial processing, which is impaired in patients with OCDS [88].

The temporal cortex is also involved in anxiety (STAI). In fact, impairments of the cortical thickness of the temporal cortex are associated with trait anxiety. The superior temporal cortex also correlates positively with trait anxiety [89, 90].

The left superior temporal cortex has been found to be linked to MFI. Studies have shown a positive relationship between activity or connectivity in the left superior temporal cortex and the severity of multidimensional fatigue symptoms. This suggests that alterations in the left superior temporal cortex may contribute to the experience of fatigue across various domains captured by the MFI.

Second, the right banks of the superior temporal sulcus was selected as a positively associated region for OCDS. The *banks of the superior temporal sulcus* is located in the superior temporal gyrus. It is involved in many processes, but this region is still not very well known. As the superior temporal region, it plays a role in speech. It is also

involved in various aspects of social perception and cognition, including recognition and interpretation of facial expressions, body language, and other non-verbal social cues. It plays a crucial role in empathy, perspective-taking, and perceiving and understanding the social interactions of others. Moreover, activity in this region has been associated with altruism and higher levels of fairness and generosity. Furthermore, this region is involved in the *theory of mind*, which is the ability to attribute mental states, beliefs, desires, and intentions. It helps the understanding and prediction of the thoughts, emotions, and behaviors of others [91, 92, 93].

While the superior temporal sulcus has not been consistently identified in the literature as a central region involved in OCDS, abnormalities in social cognition and perception have been observed in individuals with OCDS. These abnormalities may relate to the processing of social cues, emotions, and interpersonal interactions. Given the involvement of the superior temporal sulcus in social cognition, it is possible that disruptions in this region could contribute indirectly to certain aspects of OCDS.

Afterwards, the left transverse temporal white matter is present for BDI and Compulsions as a positively associated region. The *transverse temporal region*, also known as the *primary auditory cortex*, is responsible for processing auditory information. It plays a crucial role in speech understanding, tonotopic organization (mapping of sound frequency), and sound localization [94, 95].

AUD is known to impact the transverse temporal region by altering its structure and function. Moreover, studies showed that depressive disorders may also be linked with impairments of auditory processing, especially in the transverse temporal region. This can lead to difficulties in processing emotional prosody. However, the link between depression and the transverse temporal region was not extensively studied.

Regarding Compulsions, OCDS patients present a lower surface area for the transverse temporal cortex, but nothing was said for white matter [96].

Finally, the left white matter temporal pole intervenes in this thesis as a positively associated region for STAI, while the left cortex temporal pole was selected for STAI and MFI as a positively associated region as well. The *temporal poles* intervene in emotional and social cognition. It helps processing emotional information and recognizing facial expressions. Like the banks of the superior temporal sulcus, it plays a role for empathy, theory of mind, and social behavior. Moreover, it is implicated in visual processing, memory, learning, and language processing. It is also involved in decision making and value processing (evaluation of reward and punishment, integration of emotional and motivational factors to make a decision) [97].

The functional connectivity of the DMN in the temporal pole was found to be correlated with state anxiety [90].

Occipital regions

In this thesis, the left lateral occipital white matter displays both positive and negative connections for BDI and STAI. The left lateral occipital cortex also has both roles for STAI. The *lateral occipital region* is involved in object and facial recognition, as well as in motion perception [98, 99].

Parietal regions

Distinct regions within the parietal part of the brain were identified as the most recurring regions across the various behaviors.

First, the left inferior parietal white matter was selected in the five most recurring region for BDI and MFI. For MFI, it only displays positive connections, while it has both positive and negative connections for BDI. The *inferior parietal region* serves multiple functions, including spatial attention, multimodal sensory integration, visuospatial processing, oculomotor control, and perception of emotions in facial stimuli. It also plays a role in maintaining attention on the current tasks goals and responding to new stimuli in the environment. Finally, it helps understanding the perspectives of others and assessing the degree of social closeness between individuals. It involves taking the perspective of another person and overcoming one's own emotions [86, 92, 100, 101].

The white matter of the parietal lobe is known to globally be impacted by depressive disorders. It indeed plays a role in organizing, decision-making, and predicting rewards during uncertain conditions, which are essential for evaluating future response choices. Moreover, it is part of the default mode network, a network involved in introspection and self-referential thinking. Studies have observed increased cortical thickness in the left inferior parietal gyrus in depressive patients compared to healthy individuals. This finding suggests the presence structural alterations in this region in individuals with depression, supporting the results of this thesis [102].

Second, three of the superior parietal regions were selected. The *superior parietal region* is involved in visuomotor, cognitive, and sensory processing. It also plays a role for working memory, attention, and spatial orientation. Interestingly, the superior parietal region shows activation during periods of daydreaming or introspection, suggesting its involvement in self-reflection and internal mental processes. Moreover, it maintains an internal representation of the body's state, enabling individuals to be aware of their own bodies and fostering a sense of body ownership, body image, and self-representation [103, 104, 105, 106]. The left superior parietal cortex has negative associations for OCDS and Compulsions. The right superior parietal white matter intervenes as a positively associated region for BDI. Finally, the left superior parietal white matter has both positives and negatives associations for BDI, and only positive ones for MFI.

In addition to the known alterations the parietal lobe white matter in general, studies identified the bilateral superior parietal lobe as being affected by depression. These results therefore support those made in this thesis, given that the bilateral superior parietal lobe have also been identified as important for BDI [102].

Obsessive-Compulsive Disorder is also associated with structural changes in the parietal lobes [107], supporting the fact that this thesis identified the left superior parietal cortex as one of the highly predictive regions for OCDS and Compulsions.

Brain stem

The brain stem was selected as one of the five most recurring regions for OCDS, Obsessions, and Compulsions. It is responsible for many vital functions such as breath-

ing, blood pressure, heart rate, sleep, and consciousness. It also plays a role for visual reflexes, auditory processing, and is a relay center for sensory and motor pathways. Moreover, it is involved in pain suppression, and in reward pathways, influencing feelings of pleasure and motivation. It is also implicated in motor inhibition, i.e., the ability to suppress prepotent but inappropriate responses. A characteristic of addiction is a deficit of motor inhibition, causing difficulties in inhibiting inappropriate substance use or conduct [4, 108, 109], supporting the findings of this thesis. The brain stem is positively associated to all three behavioral measures. The latter are linked, and people experiencing obsessive-compulsive disorders often have uncontrollable thoughts, or engage in repetitive behavior. Impacts of the brain stem can thus reinforce the disorder.

Precuneus cortex

The precuneus cortex is a component of the default mode network, and plays a role in many complex functions. It is involved in memory processes, recollection and integration of information related to the perception of the environment. It also contributes to cue reactivity, mental imagery strategies, episodic memory retrieval, and affective responses to pain. Moreover, this region is implicated in the sense of self, agency, consciousness and self-awareness, particularly in introspective processes. It helps self-referential processing, enabling individuals to relate information to themselves. It is particularly engaged when reflecting on personal characteristics, beliefs, memories, or experiences [110, 111, 112].

The left precuneus cortex has negative associations for OCDS and Compulsions. It is not surprising, as cue reactivity is governed by the precuneus cortex and refers to the response of individuals when exposed to specific cues or triggers associated with a particular behavior or experience. Cue reactivity is extensively studied in the context of addiction and substance use disorders, where individuals may experience heightened cravings, arousal, or a strong desire to engage in substance use when exposed to cues related to their drug of choice.

Choroid plexus

The choroid plexus produces cerebrospinal fluid (CSF), regulates brain environment, removes waste, and provides a brain-blood barrier [113]. The right choroid plexus is within the five most recurring regions of OCDS, with only negative connections.

AUD is known to cause dysfunctions of the choroid plexus, impairing the blood-brain barrier and the blood-CSF-barrier. Moreover, AUD also leads to an increase of CSF volume[114, 115]. It is likely that the right choroid plexus was selected rather for the general damage caused by AUD than for its role in OCDS.

Hippocampus

The hippocampus regulates learning, spatial navigation, memory encoding and consolidation. It plays a role in flexible and goal-directed behaviors. It is also involved

in the formation and retrieval of relational memory, which is essential to remember arbitrary associations between objects or events. It is closely linked to flexible cognition and social behavior. Many studies showed that damage to the hippocampus can lead to maladaptive behavior. Moreover, it is linked to behavioral inhibition, and obsessional thinking [116]. This associates the hippocampus with the symptoms that may be experienced when being affected by OCDS. Furthermore, it has been shown that the hippocampus plays a role in OCDS [117]. In this thesis, the right hippocampus displays negative associations for Compulsions.

Mid posterior corpus callosum

The corpus callosum connects the two hemispheres of the brain, helping the integration and the transfer of information between them. It is involved in processing sensory, motor, and high-level cognitive signals, allowing communication and coordination between the two hemispheres [118].

The mid posterior part of the corpus callosum has only positive associations in the case of Obsessions. Individuals with OCDS (which also evaluates Obsessions) exhibit reduced sizes of posterior regions of the corpus callosum in comparison to individuals without the disorder. This finding suggests that the corpus callosum plays role in Obsessions [119].

Insular cortex

The right insular cortex has multiple roles and its functioning is not yet fully understood. It is implicated in gustatory and sensorimotor processing, autonomic regulation, perception of pain, auditory and vestibular functioning. It also has a role in risk-reward behavior, and self-awareness of the body and emotions. It was showed that the insular cortex is implicated in addictive habits, and damage to this region may contribute to the development of OCDS symptoms, as well as lead to difficulties recognizing both unpleasant and pleasant emotions. Furthermore, the insula is involved in emotional experiences, subjective feelings, empathy, and social cognition. Finally, it plays a role in risky decision-making and cognitive functions such as attention and salience processing since it is a component of the salience network [120, 121].

The right insular cortex is a positively associated region for Obsessions. Patients with OCDS demonstrate disrupted functioning of the insula. Imbalances in the volume of the anterior and posterior regions of the insular cortex is displayed, highlighting the contribution of the insular cortex to different symptoms of OCDS [122, 123]. It may thus have a more specific impact regarding Obsessions.

Thalamus

The thalamus plays a role in the relay and regulation of sensory and motor information. It is described as the gateway to the cerebral cortex, since it receives sensory and motor signals from the body and selectively relays and filters this information to the appropriate areas of the cortex. One of its roles is regulating arousal and attention. It helps prioritize attention by determining what sensory stimuli should receive focus

and reach conscious awareness. Moreover, it is implicated in the regulation of sleep and wakefulness. It is also responsible for the regulation of various sensory domains, except for olfaction. Furthermore, it plays a role in motor and language functions. Beyond sensory and motor functions, the thalamus is implicated in cognitive processes, such as memory, and executive functions, which encompass skills like problem-solving, decision-making, and planning. Finally, the thalamus is associated with mood and motivation. It is involved in regulating emotional states and motivational processes, contributing to the overall emotional well-being [124, 125].

The left thalamus is both a negatively and a positively associated region for Obsessions. Due to its role in attention prioritization, sensory stimulus selection, decision-making, and regulation of motivational processes, the thalamus is directly associated with obsessions. Its involvement in these cognitive and emotional functions establishes a direct connection to the manifestation of obsessive thoughts and behaviors.

Paracentral white matter

The paracentral region is involved in controlling the lower limbs. It is a part of the primary motor area [105]. The left paracentral white matter is a positively associated region within the five most recurring regions for MFI.

6.1.3 Comparison of OCDS, Obsessions, and Compulsions

Since the OCDS, Obsessions, and Compulsions scores are drawn from a same behavioral test, it is interesting to compare more in details their respective results, and how they differ from one another. The results of this thesis suggest that Obsessions do not follow the same mechanisms as OCDS and Compulsions, despite being computed from the same behavioral test.

A difference in the distribution of the selected connections between the two hemispheres can be noted. For OCDS and Compulsions, there are more connections linking the two hemispheres, while for Obsessions, the majority of the connections selected link regions of a same hemisphere.

The distribution between the different structures types linked by the edges selected, as well as the five most recurring regions selected for each of these behavioral measures, seem to support the assumption that Obsessions has different mechanisms. In fact, three of the five regions are common to OCDS and Compulsions, while the only region shared between OCDS and Obsessions, is also present in Compulsions.

For the different connections selected, OCDS and Compulsions have more negative connections than positive ones. Since the correlation obtained for the negative associations in the "Separate" model is higher than that for the positive associations, it suggests that negative connections may play a more significant role in predicting OCDS and Compulsions. On the contrary, the majority of the connections are positive for Obsessions. As the correlation obtained for the negative associations in the "Separate" model is greater than that for the positive associations, it suggests that a higher pro-

portion of positive connections is necessary to obtain significant results for Obsessions.

Finally, a notable difference between the edges selected for Obsessions, and OCDS and Compulsions, is the role of the corpus callosum. It intervenes in 25% of the selected connections for Obsessions, while only in 3% for OCDS and Compulsions.

6.2 CPM

One of the goals of this thesis was to see whether CPM, initially designed for functional data, was applicable to structural data. Shen et al. stated in their study that the results obtained using this method should lie between 0.2 and 0.5. It was also the case for the results obtained in this thesis, suggesting that CPM can be used on structural MRI data.

As said earlier, the accuracy of CPM varies on the model chosen, with the "Together" model having the least significant results. This suggests that discriminating the positive and negative associations has an important role to predict accurately. The results of the "Joint" and the "Separate" models seem to confirm this hypothesis, since the "Separate" model has five times more p-values below 0.05, seven to ten times more p-values under the Bonferroni threshold, and that the "Joint" model has respectively six and 24 times more p-values under 0.05 and the Bonferroni corrected significance threshold than the "Together" model. Moreover, these findings suggest that the best approach involves combining the discrimination between positive and negative edges during the selection process, as well as incorporating both sets of edges when constructing the model. This combined approach yields the most favorable outcomes in terms of predictive performance.

In addition to varying with the model chosen, results vary within the same model depending on the behavioral measure considered. As mentioned before, no p-value below the Bonferroni threshold were obtained for the positive associations related to BDI, Compulsions, STAI, and MFI. However, since significant results were obtained for negative associations, and positive associations in the case of OCDS and Obsessions using this model, the accuracy of its predictions is not being questioned. These findings suggest that negative connections may have a more substantial impact for the concerned behaviors, compared to positive connections, which could explain the observed results.

The consistency in the results for the "Separate" and "Together" models suggests, despite the use of linear or quadratic fits, that the choice of degree did not notably affect the statistical significance of the associations, but was more influenced by the type of model chosen.

Finally, it was shown in the results section that motion had a considerable impact on the outputs of CPM. The results obtained for BDI while using partial correlation may imply that the effects of motion have a negative predictive influence on the behavior, although accurately predicting it remains challenging. The same assumption can be drawn for OCDS, and STAI. For Compulsions and MFI, less p-values below

0.05 were found, suggesting that the effects of motion had a predictive influence on both behaviors. Regarding Obsessions, the results did not vary much when applying partial correlation, indicataing that the edges affected the most by motion were not selected to predict the behavior.

6.3 Limitations

First, some participants drank during the time of the study. It cannot be said that the abstinence was complete for every participant, influencing the changes of structural connectivity and of behavioral measures. However, it should not affect the results since abstinence itself was not entered in the model. Furthermore, the limited response rate for the third scan time has resulted in a reduced number of participants for which data was collected at t_3. Again, as we have not included time as an input of the model, it should not affect our results.

A considerable number of participants also exhibit high levels of motion, defined as an average absolute motion exceeding 0.15mm. Shen et al. suggest excluding such participants. Due to the high prevalence of the issue in our data set, this step was not taken here, potentially reducing the prediction outcomes [3]. However, the impact was reduced by using partial correlations, as described in Appendix G.

Second, the parameters used to obtain the tractograms are not optimal. As described in the methods, adjustments of the default parameters were done, improving the tractograms. Despite those, some regions had to be deleted from the connectivity matrices. The deleted regions are not known to be affected by AUD. However, a thorough optimization may lead to more accurate tractograms, allowing the use of all the connections.

Furthermore, the technique of tractography itself suffers from a limitation. The density of reconstructed connections may not accurately reflect the density of underlying white matter fibers. To address this issue, a method called *spherical-deconvolution informed filtering of tractograms (SIFT)* was developed [126, 127]. In addition to SIFT, an alternative method called SIFT2 has also been designed [128]. None of them were implemented in this thesis. Using SIFT or SIFT2 may improve biological relevance in the assessment of fiber connectivity.

At last, CPM relies on linear relationships, and may not be suitable to capture more complex or non-linear associations between connectivity and behavior. The choice of the different parameters of the CPM model influences the quality of its predictions. In this thesis, different sets of parameters were explored, some of them chosen arbitrarily (e.g. the thresholds). As it was seen in the results, the quality of the predictions varied widely. A structured comparison of the different parameters was not done, and could provide practical guidance to use CPM on structural data.

6.4 Further perspectives

First, it would be valuable to explore potential confounding factors beyond motion. Previous studies have indicated a negative correlation between the age of initial alcohol intake and the subsequent risk of developing alcohol-related disorders [4]. Additionally, research has observed differences in structural connectivity impairments between alcohol-dependent patients who underwent a single detoxification and those who experienced multiple ones [49]. Therefore, investigating the influence of the number of detoxifications on the results would be of interest. Furthermore, it would be worthwhile to investigate how relapses are linked to the obtained results. Since some of the patients relapsed during abstinence, specific biomarkers linked to this risk could possibly be identified.

As discussed in section 6.3, the SIFT method addresses the issue of tractograms not accurately reflecting the density of white matter fibers. Applying it on the tractograms to construct the connectivity matrices should have a positive impact. SIFT is designed to retrospectively enhance the accuracy of tractography results by selectively filtering out specific streamlines from the tractogram. This filtering process aims to improve the alignment between the reconstructed streamlines and the underlying diffusion images. The guidance for this filtering is provided by the outcomes of spherical deconvolution, a technique used to estimate the diffusion signal in each voxel. For SIFT2, instead of entirely removing streamlines, SIFT2 determines an appropriate cross-sectional area multiplier for each streamline. By retaining the complete streamline reconstruction while adjusting its properties, SIFT2 aims to obtain biologically meaningful measures of fiber connectivity [126, 127, 128].

Furthermore, considering the use of an alternative model, such as *Support Vector Machine regression*, may be of interest. This model is known for its resilience to overfitting and provides a weight for each edge of the connectivity matrix, eliminating the need for explicit feature selection. Additionally, employing a tool like the *Akaike Information Criterion (AIC)* can help selecting the most suitable model and parameter set. In fact, AIC identifies a model with an appropriate structure and dimension from a set of candidate models. By assessing the balance between goodness-of-fit and parsimony, the criterion determines whether a fitted model offers an optimal representation of the data while being generalizable to new data [129].

Finally, functional MRI data was also collected for the same participants. Using this data could provide additional insights, and doing a comparison between the biomarkers found using functional and structural connectomes may be interesting. Moreover, more studies cover the changes of functional connectivity in AUD, making the validation of the results easier.

Conclusion

The goal of this thesis was to identify biomarkers of behavioral recovery in patients suffering from alcohol use disorder during abstinence by studying their brain connectivity. While extensive research exists on the relationship between alcohol consumption and brain damage, the understanding of the specific connections between structural connectivity changes observed during abstinence and the corresponding changes in behavioral measures is relatively limited.

To this aim, CPM was used to predict behavioral measures from structural connectivity data, and to further select the biomarkers. At first, the raw MRI data acquired was preprocessed using the `ElikoPy` pipeline in order to correct artifacts and the effects of motion. Then labeling was done using `recon-all` to associate each voxel to a brain region, and the maps and scans were registered to a common space. Afterwards, structural connectivity matrices were built from tractograms using MSMT-CSD, and were normalized to account the effects of region volumes and fiber lengths. CPM was then used with different sets of parameters to optimize the results obtained. The best sets of parameters were chosen for each behavior. Finally, biomarkers were selected from the outputs of CPM.

The correlations between the true behavioral values and the predicted ones range from 0.252 to 0.404, which is the expected range (0.2 et 0.5) stated by Shen et al. Results show that the best approach for predicting the different behaviors combines the discrimination between positive and negative edges during the selection process, as well as incorporating both sets of edges when constructing the model.

The biomarkers selected show the global prevalence of the left hemisphere to predict the behavioral measures. As expected, they also highlight that each behavior is linked with different regions. Nonetheless, OCDS and Compulsions share a fair amount of biomarkers. In fact, the left superior parietal cortex, the brain stem, and the left precuneus cortex were identified as biomarkers for both. For OCDS, the right choroid plexus and the right banks of the superior temporal sulcus intervene as biomarkers, while the right hippocampus and the left transverse temporal white matter are biomarkers for Compulsions. BDI is mainly governed by parietal regions (right and left superior parietal white matter, and left inferior parietal white matter). The left lateral occipital white matter and the left transverse temporal white matter are also biomarkers of BDI. Regarding Obsessions, the brain stem, the mid posterior corpus callosum, the left superior temporal cortex, the right insular cortex, and the left thalamus proper are biomarkers. For STAI, temporal and occipital regions play an important role. In fact, the left superior temporal cortex, the left temporal pole (white

matter and cortex), and left lateral occipital white matter and cortex were identified as biomarkers. Finally, for MFI, biomarkers lie in the temporal and parietal regions (left superior temporal cortex, left temporal pole cortex, left inferior parietal white matter, and left superior parietal white matter). The left paracentral white matter was also identified as a biomarker. These findings are consistent with the functions of the different regions identified as biomarkers, as well as with the known regions altered by AUD or involved for the behaviors studied.

This thesis supports the fact that biomarkers of behavioral recovery in patients suffering from alcohol use disorder during abstinence can be identified. Additionally, it confirms that CPM can be used on structural data to identify consistent biomarkers. It therefore opens new horizons for a better understanding of the impacts of alcohol on structural connectivity and the links with the evolution of behavioral metrics during abstinence.

Bibliography

[1] "Understanding Alcohol Use Disorder | National Institute on Alcohol Abuse and Alcoholism (NIAAA)". (Apr. 2021), [Online]. Available: https://www.niaaa.nih.gov/publications/brochures-and-fact-sheets/understanding-alcohol-use-disorder (visited on 02/12/2023).

[2] W. H. Organization. "Alcohol Key facts". (May 2022), [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/alcohol (visited on 02/21/2023).

[3] X. Shen *et al.*, "Using connectome-based predictive modeling to predict individual behavior from brain connectivity", *Nature Protocols*, vol. 12, no. 3, pp. 506–518, Mar. 2017. DOI: 10.1038/nprot.2016.178.

[4] S. Nozaradan, J. Duque, L. Dricot, and M. Missal, *LGBIO1115 - Introduction aux neurosciences*.

[5] B. Kastler, D. Vetter, and P. Germain, *Comprendre l'IRM, Manuel d'auto-apprentissage* (Collection d'imagerie médicale diagnostic.), 7th entirely revised edition. Issy-les-Moulineaux: Elsevier Masson SAS, 2011.

[6] D. Weishaupt, V. Kochli, B. Marincek, and E. Kim, "How Does MRI Work? An Introduction to the Physics and Function of Magnetic Resonance Imaging", *Journal of Nuclear Medicine*, vol. 48, pp. 1910–1910, Nov. 2007. DOI: 10.2967/jnumed.107.045104.

[7] "Spin echo sequence - Radiology Cafe". (Mar. 2017), [Online]. Available: https://www.radiologycafe.com/frcr-physics-notes/mr-imaging/spin-echo-sequence/ (visited on 03/26/2023).

[8] "T1, T2 and PD weighted imaging - Radiology Cafe". (Mar. 2017), [Online]. Available: https://www.radiologycafe.com/frcr-physics-notes/mr-imaging/t1-t2-and-pd-weighted-imaging/ (visited on 03/26/2023).

[9] "Free interactive course on Magnetic Resonance Imaging | e-MRI". (2023), [Online]. Available: https://www.imaios.com/en/e-mri (visited on 04/03/2023).

[10] W. Van Hecke, L. Emsell, and S. Sunaert, Eds., *Diffusion Tensor Imaging: A Practical Handbook*. New York, NY: Springer, 2016. DOI: 10.1007/978-1-4939-3118-7.

[11] T. Huisman, "Diffusion-weighted and diffusion tensor imaging of the brain, made easy", *Cancer Imaging*, vol. 10, no. 1A, S163–S171, Oct. 2010. DOI: 10.1102/1470-7330.2010.9023.

[12] D. Le Bihan *et al.*, "Diffusion tensor imaging: Concepts and applications", *Journal of Magnetic Resonance Imaging*, vol. 13, no. 4, pp. 534–546, 2001. DOI: 10.1002/jmri.1076.

[13] L. J. O'Donnell and C.-F. Westin, "An introduction to diffusion tensor image analysis", *Neurosurgery clinics of North America*, vol. 22, no. 2, pp. 185–viii, Apr. 2011. DOI: 10.1016/j.nec.2010.12.004.

[14] G. A. Vorona and J. I. Berman, "Review of diffusion tensor imaging and its application in children", *Pediatric Radiology*, vol. 45, no. 3, pp. 375–381, Sep. 2015. DOI: 10.1007/s00247-015-3277-0.

[15] E. Garyfallidis, "Towards an accurate brain tractography", PhD Thesis, University of Cambridge, May 2012.

[16] T. Sarwar, K. Ramamohanarao, and A. Zalesky, "Mapping connectomes with diffusion MRI: Deterministic or probabilistic tractography?", *Magnetic Resonance in Medicine*, vol. 81, no. 2, pp. 1368–1384, Feb. 2019. DOI: 10.1002/mrm.27471.

[17] "Probabilistic tractography Introduction to dMRI". (Mar. 2023), [Online]. Available: https://carpentries-incubator.github.io/SDC-BIDS-dMRI/probabilistic_tractography/index.html (visited on 04/04/2023).

[18] Y. He and A. Evans, "Graph theoretical modeling of brain connectivity", *Current Opinion in Neurology*, vol. 23, no. 4, pp. 341–350, Aug. 2010. DOI: 10.1097/WCO.0b013e32833aa567.

[19] R. J. Wilson, *Introduction to graph theory*, 4. ed. Harlow Munich: Prentice Hall, 2009.

[20] V. Blondel and J.-C. Delvenne, *LINMA1691 - Mathématiques discrètes I : Théorie et algorithmique des graphes*.

[21] "Connectome cérébral humain Groupe d'imagerie neurofonctionnelle (GIN-IMN)". (2023), [Online]. Available: https://www.gin.cnrs.fr/fr/recherche/axe3/ (visited on 04/07/2023).

[22] J. delEtoile and H. Adeli, "Graph Theory and Brain Connectivity in Alzheimer's Disease", *The Neuroscientist*, vol. 23, no. 6, pp. 616–626, Dec. 2017. DOI: 10.1177/1073858417702621.

[23] M. Xu, *Understanding graph embedding methods and their applications*. Dec. 2020.

[24] A. koch *et al.*, "Human brain structural connectivity matricesready for modelling", *Scientific Data*, vol. 9, no. 1, p. 486, Aug. 2022. DOI: 10.1038/s41597-022-01596-9.

[25] H. Ritchie and M. Roser. "Alcohol Consumption". (Apr. 2018), [Online]. Available: https://ourworldindata.org/alcohol-consumption (visited on 02/12/2023).

[26] S. Alfonso-Loeches and C. Guerri, "Molecular and behavioral aspects of the actions of alcohol on the adult and developing brain", *Critical Reviews in Clinical Laboratory Sciences*, vol. 48, no. 1, pp. 19–47, Feb. 2011. DOI: 10.3109/10408363.2011.580567.

[27] D. Nutt *et al.*, "Alcohol and the Brain", *Nutrients*, vol. 13, no. 11, p. 3938, Nov. 2021. DOI: 10.3390/nu13113938.

[28] J. P. Connor, P. S. Haber, and W. D. Hall, "Alcohol use disorders", *The Lancet*, vol. 387, no. 10022, pp. 988–998, Mar. 2016. DOI: https://doi.org/10.1016/S0140-6736(15)00122-1.

[29] R. K. McHugh and R. D. Weiss, "Alcohol Use Disorder and Depressive Disorders", *Alcohol Research : Current Reviews*, vol. 40, no. 1, arcr.v40.1.01, Oct. 2019. DOI: 10.35946/arcr.v40.1.01.

[30] "Alcohol Effects on the Brain | Hazelden Betty Ford". (May 2015), [Online]. Available: https://www.hazeldenbettyford.org/research-studies/addiction-research/alcohol-effects-brain (visited on 04/18/2023).

[31] E. V. Sullivan, R. A. Harris, and A. Pfefferbaum, "Alcohol's Effects on Brain and Behavior", *Alcohol Research & Health*, vol. 33, no. 1-2, pp. 127–143, 2010.

[32] E. M. Müller-Oehring, Y.-C. Jung, A. Pfefferbaum, E. V. Sullivan, and T. Schulte, "The Resting Brain of Alcoholics", *Cerebral Cortex*, vol. 25, no. 11, pp. 4155–4168, Nov. 2015. DOI: 10.1093/cercor/bhu134.

[33] M. Dupuy and S. Chanraud, "Chapter One - Imaging the Addicted Brain: Alcohol", in *Imaging the Addicted Brain*, ser. International Review of Neurobiology, vol. 129, Academic Press, 2016, pp. 1–31. DOI: https://doi.org/10.1016/bs.irn.2016.04.003.

[34] S. Chanraud, A.-L. Pitel, A. Pfefferbaum, and E. V. Sullivan, "Disruption of functional connectivity of the default-mode network in alcoholism", *Cerebral Cortex (New York, N.Y.: 1991)*, vol. 21, no. 10, pp. 2272–2281, Oct. 2011. DOI: 10.1093/cercor/bhq297.

[35] "The Default Mode Network (DMN) and training its abilities in the brain". (Jul. 2021), [Online]. Available: https://www.nctneurofeedback.com/post/the-default-mode-network-dmn-and-training-its-abilities-in-the-brain (visited on 05/14/2023).

[36] K.-k. Shen *et al.*, "Structural core of the executive control network: A high angular resolution diffusion MRI study", *Human Brain Mapping*, vol. 41, no. 5, pp. 1226–1236, Nov. 2019. DOI: 10.1002/hbm.24870.

[37] G. Fein and V. A. Cardenas, "Neuroplasticity in Human Alcoholism: Studies of Extended Abstinence with Potential Treatment Implications", *Alcohol Research : Current Reviews*, vol. 37, no. 1, pp. 125–141, Jan. 2015.

[38] B. J. Weiland *et al.*, "Reduced left executive control network functional connectivity is associated with alcohol use disorders", *Alcoholism, Clinical and Experimental Research*, vol. 38, no. 9, pp. 2445–2453, Sep. 2014. DOI: 10.1111/acer.12505.

[39] W. W. Seeley, "The Salience Network: A Neural System for Perceiving and Responding to Homeostatic Demands", *Journal of Neuroscience*, vol. 39, no. 50, pp. 9878–9882, Dec. 2019. DOI: 10.1523/JNEUROSCI.1138-17.2019.

[40] E. F. Pace-Schott and D. Picchioni, "Chapter 51 - Neurobiology of Dreaming", in *Principles and Practice of Sleep Medicine (Sixth Edition)*, M. Kryger, T. Roth, and W. C. Dement, Eds., Elsevier, Jan. 2017, 529–538.e6. DOI: 10.1016/B978-0-323-24288-2.00051-9.

[41] R. N. Spreng, L. Shoemaker, and G. R. Turner, "Chapter 8 - Executive Functions and Neurocognitive Aging", in *Executive Functions in Health and Disease*, E. Goldberg, Ed., San Diego: Academic Press, Jan. 2017, pp. 169–196. DOI: 10.1016/B978-0-12-803676-1.00008-8.

[42] "Brain Basics: Know Your Brain". (Mar. 2023), [Online]. Available: https://www.ninds.nih.gov/health-information/public-education/brain-basics/brain-basics-know-your-brain (visited on 05/26/2023).

[43] "Somatosensory Systems (Section 2, Chapter 2) Neuroscience Online: An Electronic Textbook for the Neurosciences | Department of Neurobiology and Anatomy - The University of Texas Medical School at Houston". (Oct. 2020), [Online]. Available: https://nba.uth.tmc.edu/neuroscience/s2/chapter02.html (visited on 05/26/2023).

[44] "Cerebellum: Anatomy, function, and disorders". (Aug. 2022), [Online]. Available: https://www.medicalnewstoday.com/articles/313265 (visited on 05/15/2023).

[45] "Frontal Lobe Stroke: Causes, Secondary Effects, & Recovery". (Jan. 2023), [Online]. Available: https://www.flintrehab.com/frontal-lobe-stroke/ (visited on 05/15/2023).

[46] "Cerebellum Brain Damage: Causes, Effects, and Rehabilitation". (Mar. 2022), [Online]. Available: https://www.flintrehab.com/cerebellum-brain-damage/ (visited on 05/15/2023).

[47] D. Arnone, M. T. Abou-Saleh, and T. R. Barrick, "Diffusion tensor imaging of the corpus callosum in addiction", *Neuropsychobiology*, vol. 54, no. 2, pp. 107–113, 2006. DOI: 10.1159/000096992.

[48] T. Schulte, E. V. Sullivan, E. M. Müller-Oehring, E. Adalsteinsson, and A. Pfefferbaum, "Corpus callosal microstructural integrity influences interhemispheric processing: A diffusion tensor imaging study", *Cerebral Cortex (New York, N.Y.: 1991)*, vol. 15, no. 9, pp. 1384–1392, Sep. 2005. DOI: 10.1093/cercor/bhi020.

[49] O. G. O'Daly *et al.*, "Withdrawal-Associated Increases and Decreases in Functional Neural Connectivity Associated with Altered Emotional Regulation in Alcoholism", *Neuropsychopharmacology*, vol. 37, no. 10, pp. 2267–2276, Sep. 2012. DOI: 10.1038/npp.2012.77.

[50] C. Galandra, G. Basso, S. Cappa, and N. Canessa, "The alcoholic brain: Neural bases of impaired reward-based decision-making in alcohol use disorders", *Neurological Sciences*, vol. 39, no. 3, pp. 423–435, Mar. 2018. DOI: 10.1007/s10072-017-3205-1.

[51] M. Dausort and F. Falyse, "Identifying the microstructural changes for alcoholic patients during short-term abstinence via diffusion Magnetic Resonance Imaging", M.S. thesis, Ecole polytechnique de Louvain, Université catholique de Louvain, Louvain-la-Neuve, 2022.

[52] "Beck Depression Inventory (BDI)". (2018), [Online]. Available: https://arc.psych.wisc.edu/self-report/beck-depression-inventory-bdi/ (visited on 05/13/2023).

[53] *Beck's Depression Inventory Test*.

[54] R. F. Anton, D. H. Moak, and P. K. Latham, "The obsessive compulsive drinking scale: A new method of assessing outcome in alcoholism treatment studies", *Archives of General Psychiatry*, vol. 53, no. 3, pp. 225–231, Mar. 1996. DOI: 10.1001/archpsyc.1996.01830030047008.

[55] "Obsessive Compulsive Drinking Scale (OCDS)". (2018), [Online]. Available: https://arc.psych.wisc.edu/self-report/obsessive-compulsive-drinking-scale-ocds/ (visited on 05/13/2023).

[56] "The State-Trait Anxiety Inventory (STAI)". (2011), [Online]. Available: https://www.apa.org/pi/about/publications/caregivers/practice-settings/assessment/tools/trait-state (visited on 05/13/2023).

[57] E. M. Smets, B. Garssen, B. Bonke, and J. C. De Haes, "The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue", *Journal of Psychosomatic Research*, vol. 39, no. 3, pp. 315–325, Apr. 1995. DOI: 10.1016/0022-3999(94)00125-o.

[58] "Welcome to ElikoPy's documentation! ElikoPy documentation". (2021), [Online]. Available: https://elikopy.readthedocs.io/en/latest/index.html (visited on 04/07/2023).

[59] B. Jeurissen, J.-D. Tournier, T. Dhollander, A. Connelly, and J. Sijbers, "Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data", *NeuroImage*, vol. 103, pp. 411–426, Dec. 2014. DOI: 10.1016/j.neuroimage.2014.07.061.

[60] "Multi-shell multi-tissue constrained spherical deconvolution MRtrix 3.0 documentation". (Apr. 2016), [Online]. Available: https://mrtrix.readthedocs.io/en/dev/constrained_spherical_deconvolution/multi_shell_multi_tissue_csd.html (visited on 05/20/2023).

[61] "FreeSurfer Tutorial #1: Basic Terms Andy's Brain Book 1.0 documentation". (2019), [Online]. Available: https://andysbrainbook.readthedocs.io/en/latest/FreeSurfer/FS_ShortCourse/FS_01_BasicTerms.html (visited on 05/19/2023).

[62] "FreeSurfer". (Sep. 2022), [Online]. Available: https://surfer.nmr.mgh.harvard.edu (visited on 05/19/2023).

[63] "Recon-all - Free Surfer Wiki". (Dec. 2017), [Online]. Available: https://surfer.nmr.mgh.harvard.edu/fswiki/recon-all (visited on 05/19/2023).

[64] "FreeSurfer Tutorial #3: Recon-all Andy's Brain Book 1.0 documentation". (2019), [Online]. Available: https://andysbrainbook.readthedocs.io/en/latest/FreeSurfer/FS_ShortCourse/FS_03_ReconAll.html (visited on 05/19/2023).

[65] L. G. Brown, "A survey of image registration techniques", *ACM Computing Surveys*, vol. 24, no. 4, pp. 325–376, Dec. 1992. DOI: 10.1145/146370.146374.

[66] "Medical Image Registration - MATLAB & Simulink". (2023), [Online]. Available: https://www.mathworks.com/help/medical-imaging/ug/medical-image-registration.html (visited on 05/19/2023).

[67] "ElikoPy". (Jul. 2022), [Online]. Available: https://github.com/PiLAB-Medical-Imaging/elikopy (visited on 05/20/2023).

[68] "DIPY : Docs 1.7.0 - Connectivity Matrices, ROI Intersections and Density Maps". (2023), [Online]. Available: https://dipy.org/ (visited on 05/20/2023).

[69] B. Q. Rosen and E. Halgren, "A Whole-Cortex Probabilistic Diffusion Tractography Connectome", *eNeuro*, vol. 8, no. 1, ENEURO.0416–20.2020, Feb. 2021. DOI: 10.1523/ENEURO.0416-20.2020.

[70] L. Cammoun *et al.*, "Mapping the human connectome at multiple scales with diffusion spectrum MRI", *Journal of Neuroscience Methods*, vol. 203, no. 2, pp. 386–397, Jan. 2012. DOI: 10.1016/j.jneumeth.2011.09.031.

[71] Z. Gu, K. W. Jamison, M. R. Sabuncu, and A. Kuceyeski, "Heritability and interindividual variability of regional structure-function coupling", *Nature Communications*, vol. 12, no. 1, p. 4894, Aug. 2021. DOI: 10.1038/s41467-021-25184-4.

[72] A. Messé, "Parcellation influence on the connectivity-based structurefunction relationship in the human brain", *Human Brain Mapping*, vol. 41, no. 5, pp. 1167–1180, 2020. DOI: 10.1002/hbm.24866.

[73] S. Misra, O. Osogba, and M. Powers, "Chapter 1 - Unsupervised outlier detection techniques for well logs and geophysical data", in *Machine Learning for Subsurface Characterization*, S. Misra, H. Li, and J. He, Eds., Gulf Professional Publishing, Jan. 2020, pp. 1–37. DOI: 10.1016/B978-0-12-817736-5.00001-6.

[74] A. Yendiki, K. Koldewyn, S. Kakunoori, N. Kanwisher, and B. Fischl, "Spurious group differences due to head motion in a diffusion MRI study", *NeuroImage*, vol. 88, pp. 79–90, Mar. 2014. DOI: 10.1016/j.neuroimage.2013.11.027.

[75] A. Moureaux and M. Verleysen, *LGBIO2020 - Bioinstrumentation*, Louvain-la-neuve, 2021.

[76] *LINMA2472- Alogorithms in data science*, Louvain-la-Neuve, 2021.

[77] D. Hainaut and J. Laurent, *LEPL1109 - Mathématiques discrètes et probabilités*, Louvain-la-Neuve, 2020.

[78] P. Nellihela. "What is K-fold Cross Validation?" (Nov. 2022), [Online]. Available: https://towardsdatascience.com/what-is-k-fold-cross-validation-5a7bb241d82f (visited on 05/24/2023).

[79] "Numpy.polyfit NumPy v1.24 Manual". (Dec. 2022), [Online]. Available: https://numpy.org/doc/stable/reference/generated/numpy.polyfit.html (visited on 05/27/2023).

[80] "Numpy.linalg.lstsq NumPy v1.24 Manual". (Dec. 2022), [Online]. Available: https://numpy.org/doc/stable/reference/generated/numpy.linalg.lstsq.html (visited on 05/27/2023).

[81] R. A. Armstrong, "When to use the Bonferroni correction", *Ophthalmic & Physiological Optics: The Journal of the British College of Ophthalmic Opticians (Optometrists)*, vol. 34, no. 5, pp. 502–508, Sep. 2014. DOI: 10.1111/opo.12131.

[82] J. McDonald, *Multiple comparisons - Handbook of Biological Statistics*, 3rd edition. Baltimore, Maryland: Sparky House Publishing, 2014.

[83] J. Zevin, "Word Recognition", in *Encyclopedia of Neuroscience*, L. R. Squire, Ed., Oxford: Academic Press, Jan. 2009, pp. 517–522. DOI: 10.1016/B978-008045046-9.01881-7.

[84] S. J. Brooks and D. J. Stein, "Chapter 54 - Brain Volumes in Adolescents with Alcohol Use Disorder", in *Neuropathology of Drug Addictions and Substance Misuse*, V. R. Preedy, Ed., San Diego: Academic Press, Jan. 2016, pp. 587–599. DOI: 10.1016/B978-0-12-800213-1.00054-7.

[85] H.-O. Karnath, "New insights into the functions of the superior temporal cortex", *Nature Reviews Neuroscience*, vol. 2, no. 8, pp. 568–576, Aug. 2001. DOI: 10.1038/35086057.

[86] J. Radua *et al.*, "Neural response to specific components of fearful faces in healthy and schizophrenic adults", Sep. 2009. DOI: 10.1016/j.neuroimage.2009.08.030.

[87] E. D. Bigler *et al.*, "Superior temporal gyrus, language function, and autism", *Developmental Neuropsychology*, vol. 31, no. 2, pp. 217–238, 2007. DOI: 10.1080/87565640701190841.

[88] J.-S. Choi *et al.*, "Morphometric alterations of anterior superior temporal cortex in obsessive-compulsive disorder", *Depression and Anxiety*, vol. 23, no. 5, pp. 290–296, 2006. DOI: 10.1002/da.20171.

[89] F. Saviola, E. Pappaianni, A. Monti, A. Grecucci, J. Jovicich, and N. De Pisapia, "Trait and state anxiety are mapped differently in the human brain", *Scientific Reports*, vol. 10, no. 1, p. 11 112, Jul. 2020. DOI: 10.1038/s41598-020-68008-z.

[90] Y. Li and L. Jiang, "State and Trait Anxiety Share Common Network Topological Mechanisms of Human Brain", *Frontiers in Neuroinformatics*, vol. 16, p. 859 309, Jun. 2022. DOI: 10.3389/fninf.2022.859309.

[91] F. Leroy *et al.*, "New human-specific brain landmark: The depth asymmetry of superior temporal sulcus", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 4, pp. 1208–1213, Jan. 2015. DOI: 10.1073/pnas.1412389112.

[92] V. E. Sturm, C. M. Haase, and R. W. Levenson, "Chapter 22 - Emotional Dysfunction in Psychopathology and Neuropathology: Neural and Genetic Pathways", in *Genomics, Circuits, and Pathways in Clinical Neuropsychiatry*, T. Lehner, B. L. Miller, and M. W. State, Eds., San Diego: Academic Press, Jan. 2016, pp. 345–364. DOI: 10.1016/B978-0-12-800105-9.00022-6.

[93] H. Bukowski and C. Lamm, "Superior Temporal Sulcus", in *Encyclopedia of Personality and Individual Differences*, V. Zeigler-Hill and T. K. Shackelford, Eds., Cham: Springer International Publishing, 2018, pp. 1–5. DOI: 10.1007/978-3-319-28099-8_463-1.

[94] J. I. Skipper and S. L. Small, "fMRI Studies of Language", in *Encyclopedia of Language & Linguistics (Second Edition)*, K. Brown, Ed., Oxford: Elsevier, Jan. 2006, pp. 496–511. DOI: 10.1016/B0-08-044854-2/02399-3.

[95] "Transverse temporal gyri - e-Anatomy - IMAIOS". (2023), [Online]. Available: https://www.imaios.com/en/e-anatomy/anatomical-structure/transverse-temporal-gyri-1553798716 (visited on 05/30/2023).

[96] P. S. Boedhoe *et al.*, "Cortical Abnormalities Associated With Pediatric and Adult Obsessive-Compulsive Disorder: Findings From the ENIGMA Obsessive-Compulsive Disorder Working Group", *American Journal of Psychiatry*, vol. 175, no. 5, pp. 453–462, May 2018. DOI: 10.1176/appi.ajp.2017.17050485.

[97] B. Herlin, V. Navarro, and S. Dupont, "The temporal pole: From anatomy to functionA literature appraisal", *Journal of Chemical Neuroanatomy*, vol. 113, p. 101 925, Apr. 2021. DOI: 10.1016/j.jchemneu.2021.101925.

[98] K. Grill-Spector, Z. Kourtzi, and N. Kanwisher, "The lateral occipital complex and its role in object recognition", *Vision Research*, vol. 41, no. 10-11, pp. 1409–1422, 2001. DOI: 10.1016/s0042-6989(01)00073-6.

[99] A. H. Palejwala *et al.*, "Anatomy and white matter connections of the lateral occipital cortex", *Surgical and radiologic anatomy: SRA*, vol. 42, no. 3, pp. 315–328, Mar. 2020. DOI: 10.1007/s00276-019-02371-z.

[100] V. Singh-Curry and M. Husain, "The functional role of the inferior parietal lobe in the dorsal and ventral stream dichotomy", *Neuropsychologia*, vol. 47, no. 6, pp. 1434–1448, May 2009. DOI: 10.1016/j.neuropsychologia.2008.11.033.

[101] D. M. Clower, R. A. West, J. C. Lynch, and P. L. Strick, "The Inferior Parietal Lobule Is the Target of Output from the Superior Colliculus, Hippocampus, and Cerebellum", *The Journal of Neuroscience*, vol. 21, no. 16, pp. 6283–6291, Aug. 2001. DOI: 10.1523/JNEUROSCI.21-16-06283.2001.

[102] F.-F. Zhang, W. Peng, J. A. Sweeney, Z.-Y. Jia, and Q.-Y. Gong, "Brain structure alterations in depression: Psychoradiological evidence", *CNS Neuroscience & Therapeutics*, vol. 24, no. 11, pp. 994–1003, Mar. 2018. DOI: 10.1111/cns.12835.

[103] D. M. Wolpert, S. J. Goodbody, and M. Husain, "Maintaining internal representations: The role of the human superior parietal lobe", *Nature Neuroscience*, vol. 1, no. 6, pp. 529–533, Oct. 1998. DOI: 10.1038/2245.

[104] M. Koenigs, A. K. Barbey, B. R. Postle, and J. Grafman, "Superior Parietal Cortex Is Critical for the Manipulation of Information in Working Memory", *The Journal of Neuroscience*, vol. 29, no. 47, pp. 14 980–14 986, Nov. 2009. DOI: 10.1523/JNEUROSCI.3706-09.2009.

[105] P. Johns, "Chapter 3 - Functional neuroanatomy", in *Clinical Neuroscience*, P. Johns, Ed., Churchill Livingstone, Jan. 2014, pp. 27–47. DOI: 10.1016/B978-0-443-10321-6.00003-5.

[106] A. A. S. Alahmadi, "Investigating the sub-regions of the superior parietal cortex using functional magnetic resonance imaging connectivity", *Insights into Imaging*, vol. 12, no. 1, p. 47, Apr. 2021. DOI: 10.1186/s13244-021-00993-9.

[107] F. Piras, F. Piras, C. Chiapponi, P. Girardi, C. Caltagirone, and G. Spalletta, "Widespread structural brain changes in OCD: A systematic review of voxel-based morphometry studies", *Cortex*, Special issue: The clinical anatomy of the limbic lobe and connected structures, vol. 62, pp. 89–108, Jan. 2015. DOI: 10.1016/j.cortex.2013.01.016.

[108] "Brainstem: Overview, Function & Anatomy". (Jun. 2021), [Online]. Available: https://my.clevelandclinic.org/health/body/21598-brainstem (visited on 05/31/2023).

[109] H. Basinger and J. P. Hogg, "Neuroanatomy, Brainstem", in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2023.

[110] M.-S. Lee, "Chapter 26 - Methods of neuroimaging in depression: Applications to resting-state functional connectivity", in *The Neuroscience of Depression*, C. R. Martin, L.-A. Hunter, V. B. Patel, V. R. Preedy, and R. Rajendram, Eds., Academic Press, Jan. 2021, pp. 261–270. DOI: 10.1016/B978-0-12-817933-8.00091-8.

[111] E. T. Rolls, "Chapter 1 - The neuroscience of emotional disorders", in *Handbook of Clinical Neurology*, ser. Disorders of Emotion in Neurologic Disease, K. M. Heilman and S. E. Nadeau, Eds., vol. 183, Elsevier, Jan. 2021, pp. 1–26. DOI: 10.1016/B978-0-12-822290-4.00002-5.

[112] D. Borsook, N. Maleki, and R. Burstein, "Chapter 42 - Migraine", in *Neurobiology of Brain Disorders*, M. J. Zigmond, L. P. Rowland, and J. T. Coyle, Eds., San Diego: Academic Press, Jan. 2015, pp. 693–708. DOI: 10.1016/B978-0-12-398270-4.00042-2.

[113] K. Javed, V. Reddy, and F. Lui, "Neuroanatomy, Choroid Plexus", in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2023.

[114] P. F. Nixon, L. Jordan, C. Zimitat, S. E. Rose, and F. Zelaya, "Choroid plexus dysfunction: The initial event in the pathogenesis of Wernicke's encephalopathy and ethanol intoxication", *Alcoholism, Clinical and Experimental Research*, vol. 32, no. 8, pp. 1513–1523, Aug. 2008. DOI: 10.1111/j.1530-0277.2008.00723.x.

[115] A. Pfefferbaum, E. V. Sullivan, M. J. Rosenbloom, P. K. Shear, D. H. Mathalon, and K. O. Lim, "Increase in brain cerebrospinal fluid volume is greater in older than in younger alcoholic patients: A replication study and CT/MRI comparison", *Psychiatry Research: Neuroimaging*, vol. 50, no. 4, pp. 257–274, Dec. 1993. DOI: 10.1016/0925-4927(93)90004-2.

[116] "Hippocampus Functions". (Aug. 2019), [Online]. Available: https://www.news-medical.net/health/Hippocampus-Functions.aspx (visited on 05/31/2023).

[117] T. J. Reess *et al.*, "Association between hippocampus volume and symptom profiles in obsessivecompulsive disorder", *NeuroImage : Clinical*, vol. 17, pp. 474–480, Nov. 2017. DOI: 10.1016/j.nicl.2017.11.006.

[118] A. Goldstein, B. P. Covington, N. Mahabadi, and F. B. Mesfin, "Neuroanatomy, Corpus Callosum", in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2023.

[119] M. Di Paola *et al.*, "The structure of the corpus callosum in obsessive compulsive disorder", *European Psychiatry: The Journal of the Association of European Psychiatrists*, vol. 28, no. 8, pp. 499–506, Oct. 2013. DOI: 10.1016/j.eurpsy.2012.07.001.

[120] "Insula". (2021), [Online]. Available: https://www.physio-pedia.com/Insula (visited on 05/31/2023).

[121] L. Q. Uddin, J. S. Nomi, B. Hebert-Seropian, J. Ghaziri, and O. Boucher, "Structure and function of the human insula", *Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society*, vol. 34, no. 4, pp. 300–306, Jul. 2017. DOI: 10.1097/WNP.0000000000000377.

[122] G. K. Eng *et al.*, "Relationships between interoceptive sensibility and resting-state functional connectivity of the insula in obsessivecompulsive disorder", *Cerebral Cortex*, vol. 32, no. 23, pp. 5285–5300, Dec. 2022. DOI: 10.1093/cercor/bhac014.

[123] A. Song *et al.*, "Disproportionate Alterations in the Anterior and Posterior Insular Cortices in ObsessiveCompulsive Disorder", *PLoS ONE*, vol. 6, no. 7, e22361, Jul. 2011. DOI: 10.1371/journal.pone.0022361.

[124] "Thalamus: What It Is, Function & Disorders". (Mar. 2022), [Online]. Available: https://my.clevelandclinic.org/health/body/22652-thalamus (visited on 05/31/2023).

[125] T. J. Torrico and S. Munakomi, "Neuroanatomy, Thalamus", in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2023.

[126] R. E. Smith, J.-D. Tournier, F. Calamante, and A. Connelly, "SIFT: Spherical-deconvolution informed filtering of tractograms", *NeuroImage*, vol. 67, pp. 298–312, Feb. 2013. DOI: 10.1016/j.neuroimage.2012.11.049.

[127] R. E. Smith, J.-D. Tournier, F. Calamante, and A. Connelly, "The effects of SIFT on the reproducibility and biological accuracy of the structural connectome", *NeuroImage*, vol. 104, pp. 253–265, Jan. 2015. DOI: 10.1016/j.neuroimage.2014.10.004.

[128] R. E. Smith, J.-D. Tournier, F. Calamante, and A. Connelly, "SIFT2: Enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography", *NeuroImage*, vol. 119, pp. 338–351, Oct. 2015. DOI: 10.1016/j.neuroimage.2015.06.092.

[129] J. E. Cavanaugh and A. A. Neath, "The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements", *WIREs Computational Statistics*, vol. 11, no. 3, e1460, May 2019. DOI: 10.1002/wics.1460.

[130] Q. Dessain and M. Simon, "Development of an image processing pipeline for large-scale in vivo population studies of the brain microstructure", M.S. thesis, Université catholique de Louvain, Louvain-la-Neuve, 2021.

Appendices

Appendix A

ElikoPy

`ElikoPy` is a `Python` library that facilitates the processing of diffusion imaging for microstructural analysis. It is a preprocessing pipeline that can correct diffusion MRI artifacts and provide information on the brain microstructure. Moreover, it provides a full quality assessment and quality reports for each processed subject [58, 130].

It is based on the following existing `Python` libraries: `Dipy`, Microstructure fingerprinting (MF) and `Dmipy`. It also uses FSL, a library providing analysis tools for brain imaging data such as MRI and DTI [58, 130].

The different steps of the `ElikoPy` pipeline can be seen on Fig. A.1. The first step of the pipeline is to check whether the dimensions of the different files are correct. If it is the case, a storage folder is created for each subject. It will contain the results of each of the steps applied to the subject. Preprocessing can then be applied. It includes different optional steps, such as reslicing, brain extraction, Principal Component Analysis (PCA) based denoising using the Marchenko-Pastur distribution, suppression of Gibbs ringing artifacts, estimation and correction of the susceptibility induced and Eddy current induced distortions, modeling and correction of subject movements, and finally bias field correction [58, 130].

A white matter mask is then obtained. Afterwards, different models, such as DTI, can be applied to do a microstructure analysis. The metrics obtained from the chosen models can be registered to a common space, and statistics can be computed [58, 130].

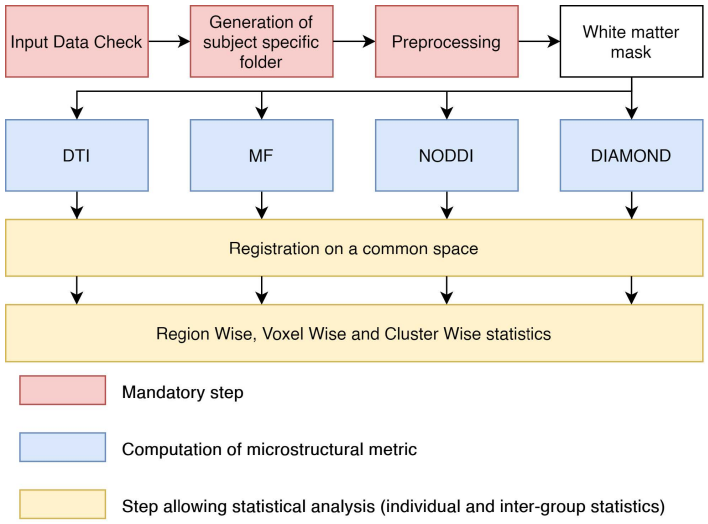


Figure A.1: ElikoPy pipeline steps. From [58].

Appendix B

Number of labels by map produced by recon-all

Table B.1: Number of labels of each parcellation maps produced by the `recon-all` *command.*

Map	Number of labels
wmparc	177
aparc+aseg	110
aparc.DKTatlas+aseg	104
aseg.auto	44
aseg	44
aseg.presurf.hypos	44
aseg.presurf	44
aseg.auto_noCCseg	39
ribbon	5
filled.auto	3
filled	3

Appendix C

Boxplots used to detect the outliers for the different behavior metrics

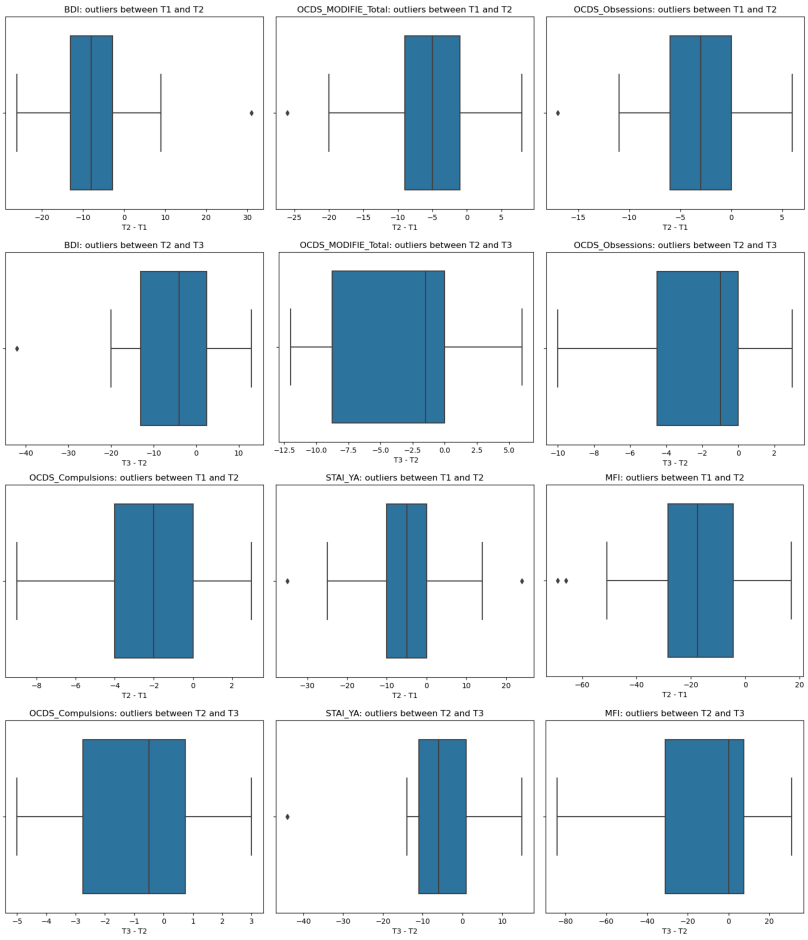


Figure C.1: Boxplots from which the outliers were identified.

Appendix D

Different sets of parameters used for CPM

Table D.1: Different sets of parameters used for CPM.

Parameters			
Model	Correlation	Order of the polynomial fit	Threshold
Separate	Pearson	1	0.01 0.001 0.05 0.1
		2	0.01 0.001 0.05 0.1
	Spearman	1	0.01 0.001 0.05 0.1
		2	0.01 0.001 0.05 0.1
Joint	Pearson	1	0.01 0.001 0.05 0.1
	Spearman	1	0.01 0.001 0.05 0.1

Table D.2: Continuation of table D.1: Different sets of parameters used for CPM.

Parameters			
Model	Correlation	Order of the polynomial fit	Threshold
Together	Pearson	1	0.01 0.001 0.05 0.1
		2	0.01 0.001 0.05 0.1
	Spearman	1	0.01 0.001 0.05 0.1
		2	0.01 0.001 0.05 0.1

Appendix E

Results using partial correlation

E.1 Correlations

E.1.1 "Separate" Model – Positive edges

Table E.1: Correlations for the different sets of parameters for the "Separate" model using partial correlation. Only the positive associations are represented in the table. The correlations highlighted in blue have a p-value below the significance threshold of 0.05. The correlations highlighted in orange have a p-value below the Bonferroni modified threshold of 0.00125. The correlations highlighted in red have the best p-value across the whole table, with a p-value below the Bonferroni modified threshold. The negative correlation were not highlighted even if they have a significant p-value since they do not predict good results.

Parameters			Correlation for the different behaviors					
Correlation	Order	Threshold	BDI	OCDS	Obs.	Comp.	STAI	MFI
Pearson	1	0.01	0.192	0.287	0.329	0.209	0.173	0.216
		0.001	0.031	0.094	0.240	0.110	not converged	0.053
		0.05	0.169	0.274	0.300	0.208	0.184	0.207
		0.1	0.171	0.285	0.303	0.224	0.254	0.239
	2	0.01	0.184	0.288	0.324	0.216	0.173	0.209
		0.001	0.043	0.118	0.248	0.117	not converged	0.065
		0.05	0.168	0.267	0.289	0.208	0.181	0.199
		0.1	0.170	0.276	0.292	0.219	0.253	0.234
Spearman	1	0.01	0.260	0.154	0.303	0.201	0.083	0.226
		0.001	0.066	0.289	0.050	0.251	-0.029	0.053
		0.05	0.216	0.260	0.291	0.167	0.249	0.263
		0.1	0.208	0.275	0.296	0.194	0.241	0.254
	2	0.01	0.253	0.114	0.303	0.180	0.078	0.213
		0.001	0.069	0.277	0.081	0.244	-0.048	0.060
		0.05	0.203	0.255	0.268	0.152	0.240	0.252
		0.1	0.192	0.263	0.276	0.173	0.232	0.241

E.1.2 "Separate" Model – Negative edges

Table E.2: Correlations for the different sets of parameters for the "Separate" model using partial correlation. Only the negative associations are represented in the table. The correlations highlighted in blue have a p-value below the significance threshold of 0.05. The correlations highlighted in orange have a p-value below the Bonferroni modified threshold of 0.00125. The correlations highlighted in red have the best p-value across the whole table, with a p-value below the Bonferroni modified threshold.

Parameters			Correlation for the different behaviors					
Correlation	Order	Threshold	BDI	OCDS	Obs.	Comp.	STAI	MFI
Pearson	1	0.01	0.104	0.216	0.260	0.150	0.297	0.239
		0.001	0.221	0.330	0.345	0.404	not converged	0.343
		0.05	0.161	0.180	0.243	0.274	0.245	0.218
		0.1	0.178	0.218	0.221	0.187	0.214	0.216
	2	0.01	0.087	0.227	0.275	0.168	0.307	0.249
		0.001	0.259	0.337	0.365	0.397	not converged	0.373
		0.05	0.195	0.195	0.250	0.283	0.273	0.254
		0.1	0.203	0.229	0.216	0.195	0.226	0.251
Spearman	1	0.01	0.033	0.185	0.189	0.198	0.290	0.167
		0.001	0.240	0.320	0.241	0.328	0.211	0.292
		0.05	0.062	0.189	0.225	0.166	0.049	0.160
		0.1	0.025	0.147	0.183	0.143	0.074	0.164
	2	0.01	0.035	0.188	0.195	0.198	0.318	0.233
		0.001	0.242	0.322	0.266	0.324	0.192	0.353
		0.05	0.089	0.188	0.218	0.160	0.065	0.218
		0.1	0.047	0.150	0.177	0.142	0.082	0.218

E.1.3 "Joint" Model

Table E.3: Correlations for the different sets of parameters for the "Joint" model using partial correlation. The correlations highlighted in blue have a p-value below the significance threshold of 0.05. The correlations highlighted in orange have a p-value below the Bonferroni modified threshold of 0.00125. The correlations highlighted in red have the best p-value across the whole table, with a p-value below the Bonferroni modified threshold.

Parameters		Correlation for the different behaviors					
Correlation	Threshold	BDI	OCDS	Obs.	Comp.	STAI	MFI
Pearson	0.01	0.203	0.290	0.336	0.215	0.341	0.265
	0.001	0.146	0.309	0.363	0.344	0.208	0.260
	0.05	0.248	0.277	0.330	0.311	0.344	0.265
	0.1	0.275	0.317	0.326	0.262	0.386	0.288
Spearman	0.01	0.190	0.208	0.321	0.251	0.270	0.234
	0.001	0.214	0.369	0.206	0.362	0.158	0.145
	0.05	0.221	0.283	0.309	0.215	0.222	0.279
	0.1	0.200	0.291	0.310	0.227	0.249	0.278

E.1.4 "Together" Model

Table E.4: Correlations for the different sets of parameters for the "Together" model using partial correlation. The correlations highlighted in blue have a p-value below the significance threshold of 0.05. The correlations highlighted in orange have a p-value below the Bonferroni modified threshold of 0.00125. The correlations highlighted in red have the best p-value across the whole table, with a p-value below the Bonferroni modified threshold. The negative correlation were not highlighted even if they have a significant p-value since they do not predict good results.

Parameters			Correlation for the different behaviors					
Correlation	Order	Threshold	BDI	OCDS	Obs.	Comp.	STAI	MFI
Pearson	1	0.01	0.181	0.219	0.229	0.098	-0.021	0.057
		0.001	-0.017	-0.089	0.140	-0.429	-0.1511	-0.200
		0.05	0.126	0.175	0.104	-0.062	0.01268	0.025
		0.1	0.107	0.115	0.014	0.007	0.077	0.055
	2	0.01	0.172	0.190	0.220	0.037	0.029	0.014
		0.001	-0.010	-0.150	0.105	-0.143	-0.188	-0.253
		0.05	0.134	0.181	0.206	0.010	0.055	0.025
		0.1	0.128	0.191	0.158	0.081	0.117	0.076
Spearman	1	0.01	0.128	0.052	0.008	0.107	-0.584	0.019
		0.001	-0.561	-0.069	0.182	0.078	0.028	0.281
		0.05	0.056	-0.108	0.062	0.015	-0.333	-0.592
		0.1	0.089	-0.323	-0.130	-0.061	-0.457	-0.479
	2	0.01	0.147	0.031	-0.030	0.074	-0.174	0.071
		0.001	-0.523	0.043	0.209	0.087	0.013	0.361
		0.05	0.126	-0.138	0.009	-0.018	0.065	-0.188
		0.1	0.161	-0.271	-0.169	-0.049	0.125	-0.100

E.2 *p*-values

E.2.1 "Separate" Model – Positive edges

Table E.5: p-values for the different sets of parameters for the "Separate" model using partial correlation. Only the positive associations are represented in the table. The p-values highlighted in blue are below the significance threshold of 0.05. The p-values highlighted in orange are below the Bonferroni modified threshold of 0.00125. The p-values highlighted in red are the best p-values across the whole table, and are below the Bonferroni modified threshold. The p-values associated with a negative correlation are not highlighted even if they are significant since they do not predict good results.

Parameters			*p*-values for the different behaviors					
Correlation	Order	Threshold	BDI	OCDS	Obs.	Comp.	STAI	MFI
Pearson	1	0.01	0.025	9.5e-04	1.1e-04	0.016	0.047	0.012
		0.001	0.717	0.287	0.005	0.208	not converged	0.541
		0.05	0.049	0.002	4.9e-04	0.017	0.033	0.017
		0.1	0.046	0.001	4.2e-04	0.009	0.003	0.005
	2	0.01	0.031	0.009	1.5e-04	0.013	0.046	0.015
		0.001	0.612	0.182	0.004	0.180	not converged	0.452
		0.05	0.051	0.002	8e-04	0.016	0.036	0.021
		0.1	0.048	0.002	6.9e-04	0.011	0.003	0.006
Spearman	1	0.01	0.002	0.080	4.2e-04	0.021	0.343	0.009
		0.001	0.443	8.6e-04	0.570	0.003	0.734	0.540
		0.05	0.011	0.002	7.3e-04	0.055	0.003	0.002
		0.1	0.015	0.002	5.9e-04	0.026	0.005	0.003
	2	0.01	0.003	0.194	4.1e-04	0.103	0.373	0.013
		0.001	0.425	0.001	0.354	0.004	0.578	0.493
		0.05	0.017	0.003	0.002	0.082	0.005	0.003
		0.1	0.025	0.002	0.001	0.047	0.007	0.005

E.2.2 "Separate" Model – Negative edges

Table E.6: p-values for the different sets of parameters for the "Separate" model using partial correlation. Only the negative associations are represented in the table. The p-values highlighted in blue are below the significance threshold of 0.05. The p-values highlighted in orange are below the Bonferroni modified threshold of 0.00125. The p-values highlighted in red are the best p-values across the whole table, and are below the Bonferroni modified threshold. The p-values associated with a negative correlation are not highlighted even if they are significant since they do not predict good results.

Parameters			p-values for the different behaviors					
Correlation	Order	Thresh.	BDI	OCDS	Obs.	Comp.	STAI	MFI
Pearson	1	0.01	0.228	0.013	0.002	0.087	5.3e-04	0.005
		0.001	0.009	1.3e-04	5.36e-05	1.67e-06	not converged	5.5e-05
		0.05	0.061	0.045	0.005	0.001	0.004	0.011
		0.1	0.037	0.012	0.011	0.032	0.013	0.012
	2	0.01	0.313	0.009	0.001	0.054	3.3e-04	0.003
		0.001	0.002	9.23e-05	1.72e-05	2.62e-05	not converged	1.01e-05
		0.05	0.022	0.026	0.003	0.001	0.002	0.003
		0.1	0.017	0.008	0.013	0.025	0.008	0.003
Spearman	1	0.01	0.697	0.035	0.030	0.023	7.1e-04	0.054
		0.001	0.005	2.1e-04	0.201	1.2e-04	0.015	6.7e-04
		0.05	0.469	0.031	0.009	0.056	0.574	0.065
		0.1	0.770	0.096	0.036	0.103	0.394	0.058
	2	0.01	0.683	0.032	0.025	0.022	1.9e-04	0.007
		0.001	0.004	1.9e-04	0.002	1.5e-04	0.026	3.29e-05
		0.05	0.301	0.032	0.012	0.066	0.453	0.011
		0.1	0.583	0.089	0.042	0.104	0.344	0.011

E.2.3 "Joint" Model

Table E.7: p-values for the different sets of parameters for the "Joint" model using partial correlation. The p-values highlighted in blue are below the significance threshold of 0.05. The p-values highlighted in orange are below the Bonferroni modified threshold of 0.00125. The p-values highlighted in red are the best p-values across the whole table, and are below the Bonferroni modified threshold. The p-values associated with a negative correlation are not highlighted even if they are significant since they do not predict good results.

Parameters		p-values for the different behaviors					
Correlation	Threshold	BDI	OCDS	Obs.	Comp.	STAI	MFI
Pearson	0.01	0.017	8.4e-04	8.45e-05	0.013	6.03e-05	0.002
	0.001	0.090	3.5e-04	1.99e-05	5.55e-05	0.016	0.002
	0.05	0.003	0.002	1.1e-04	2.9e-04	5.11e-05	0.002
	0.1	0.001	2.5e-04	1.3e-04	0.002	4.69e-06	7.8e-04
Spearman	0.01	0.026	0.017	1.8e-04	0.003	0.002	0.006
	0.001	0.012	1.67e-05	0.018	2.09e-05	0.070	0.096
	0.05	0.009	0.001	3.e-041	0.013	0.010	0.001
	0.1	0.019	7.9e-04	3e-04	0.009	0.003	0.001

E.2.4 "Together" Model

Table E.8: p-values for the different sets of parameters for the "Together" model using partial correlation. The p-values highlighted in blue are below the significance threshold of 0.05. The p-values highlighted in orange are below the Bonferroni modified threshold of 0.00125. The p-values highlighted in red are the best p-values across the whole table, and are below the Bonferroni modified threshold. The p-values associated with a negative correlation are not highlighted even if they are significant since they do not predict good results.

Parameters			*p*-values for the different behaviors					
Correlation	Order	Thresh.	BDI	OCDS	Obs.	Comp.	STAI	MFI
Pearson	1	0.01	0.034	0.012	0.008	0.262	0.809	0.511
		0.001	0.838	0.314	0.110	3.08e-07	0.083	0.021
		0.05	0.142	0.046	0.233	0.481	0.885	0.767
		0.1	0.213	0.193	0.869	0.935	0.379	0.529
	2	0.01	0.045	0.030	0.011	0.672	0.74	0.864
		0.001	0.900	0.088	0.230	0.102	0.030	0.003
		0.05	0.121	0.039	0.018	0.902	0.530	0.773
		0.1	0.136	0.029	0.069	0.355	0.181	0.381
Spearman	1	0.01	0.137	0.551	0.927	0.220	1.97e-13	0.821
		0.001	1.43e-12	0.435	0.037	0.374	0.743	0.001
		0.05	0.515	0.223	0.478	0.861	9.113e-05	6.92e-14
		0.1	0.304	1.8e-04	0.136	0.483	3.52e-08	5.95e-09
	2	0.01	0.087	0.724	0.729	0.400	0.045	0.412
		0.001	7.32e-11	0.620	0.016	0.318	0.874	2.07e-05
		0.05	0.145	0.118	0.916	0.831	0.453	0.030
		0.1	0.061	0.001	0.053	0.575	0.153	0.253

Appendix F

Connections selected with a threshold of 20 edges

F.1 BDI

Table F.1: Selected connections for BDI, as well as the model in which they were selected.

Connections		**Models that selected the connections**
Mid Posterior Corpus Callosum	Right medial orbito frontal cortex	Positive associations only
Posterior Corpus Callosum	Mid Posterior Corpus Callosum	"Joint" positive associations
Left Caudate	Right Putamen	Positive associations only
Left Caudate	Left caudal middle frontal white matter	Positive associations only
Left Caudate	Right transverse temporal white matter	Positive associations only, "Joint" positive associations
Left Cerebellum Cortex	Central Corpus Callosum	"Joint" positive associations
Left Cerebellum White Matter	Left middle temporal white matter	"Joint" negative associations
Left Inferior Lateral Ventricle	Right middle temporal cortex	"Joint" negative associations
Left Putamen	Left superior parietal white matter	Positive associations only
Left Thalamus Proper	Right precuneus cortex	"Joint" negative associations, Negative associations only
Right Cerebellum Cortex	Right isthmus cingulate cortex	"Joint" negative associations
Right Cerebellum White Matter	Right cuneus cortex	"Joint" negative associations

Right Cerebellum White Matter	Right middle temporal white matter	"Joint" negative associations
Right choroid plexus	Right isthmus cingulate cortex	"Joint" negative associations
Left bankssts cortex	Left posterior cingulate white matter	Positive associations only, "Joint" positive associations
Left caudal anterior cingulate cortex	Posterior Corpus Callosum	Positive associations only, "Joint" positive associations
Left frontal pole cortex	Right lateral orbito frontal cortex	"Joint" negative associations
Left inferior temporal cortex	Right inferior parietal cortex	"Joint" negative associations
Left insula cortex	Left inferior parietal white matter	"Joint" positive associations
Left posterior cingulate cortex	Left temporal pole white matter	Positive associations only
Left precuneus cortex	Right Thalamus Proper	"Joint" negative associations, Negative associations only
Left precuneus cortex	Left pars triangularis white matter	"Joint" negative associations
Left superior frontal cortex	Left pars triangularis white matter	"Joint" negative associations
Left superior parietal cortex	Right Thalamus Proper	"Joint" negative associations
Left superior parietal cortex	Right Ventral Diencephalon	"Joint" negative associations
Right caudal middle frontal cortex	Right precentral white matter	"Joint" positive associations
Right lingual cortex	Right precuneus white matter	"Joint" positive associations
Right lingual cortex	Right superior parietal white matter	"Joint" positive associations
Right medial orbital frontal cortex	Right superior parietal white matter	Positive associations only
Right parahippocampal cortex	Right superior parietal white matter	"Joint" positive associations
Right precuneus cortex	Right superior parietal white matter	"Joint" positive associations
Right transverse temporal cortex	Right superior parietal white matter	Positive associations only
Left frontal pole white matter	Right lateral orbito frontal cortex	"Joint" negative associations
Left inferior parietal white matter	Left insula white matter	"Joint" positive associations

Left inferior parietal white matter	Left paracentral white matter	Positive associations only, "Joint" positive associations
Left lateral occipital white matter	Left superior parietal white matter	Positive associations only
Left medial orbito frontal white matter	Left transverse temporal white matter	Positive associations only
Left middle temporal white matter	Right fusiform cortex	"Joint" negative associations
Left precuneus white matter	Right Thalamus Proper	"Joint" negative associations
Left rostral middle frontal white matter	Right lateral orbito frontal cortex	"Joint" negative associations
Left rostral middle frontal white matter	Left transverse temporal white matter	"Joint" positive associations
Left superior frontal white matter	Left transverse temporal white matter	"Joint" positive associations
Left superior parietal white matter	Right Ventral Diencephalon	"Joint" negative associations
Left supra marginal white matter	Left insula white matter	"Joint" positive associations
Left temporal pole white matter	Right lateral orbito frontal cortex	"Joint" positive associations
Left transverse temporal white matter	Right rostral middle frontal white matter	Positive associations only
Right bankssts white matter	Right superior parietal white matter	Positive associations only, "Joint" positive associations
Right caudal middle frontal white matter	Right parso percularis white matter	Positive associations only
Right inferior parietal white matter	Right supra marginal white matter	Positive associations only
Right lingual white matter	Right precuneus white matter	Positive associations only, "Joint" positive associations
Right middle temporal white matter	Right parsorbitalis white matter	"Joint" negative associations
Right superior parietal white matter	Right insula white matter	Positive associations only
Right superior parietal white matter	Right superior temporal white matter	Positive associations only, "Joint" positive associations

F.2 OCDS

Table F.2: Selected connections for OCDS, as well as the model in which they were selected.

Connections		**Models that selected the connections**
Left Accumbens area	Left lingual white matter	"Joint" negative associations
Left Accumbens area	Left temporal pole white matter	"Joint" negative associations
Left Lateral Ventricle	Right medial orbito frontal white matter	"Joint" negative associations
Left Thalamus Proper	Right precuneus cortex	Negative associations only
Right Cerebellum Cortex	Right bankssts white matter	Positive associations only, "Joint" positive associations
Right Hippocampus	Right entorhinal white matter	"Joint" negative associations, Negative associations only
Right Hippocampus	Right frontal pole white matter	"Joint" negative associations
Right choroid plexus	Right superior frontal cortex	"Joint" negative associations
Right choroid plexus	Right superior frontal white matter	"Joint" negative associations, Negative associations only
Left lateral occipital cortex	Right inferior temporal cortex	Positive associations only, "Joint" positive associations
Left precuneus cortex	Right Pallidum	"Joint" negative associations
Left precuneus cortex	Right Thalamus Proper	"Joint" negative associations
Left precuneus cortex	Right inferior parietal cortex	"Joint" negative associations
Left precuneus cortex	Right medial orbito frontal white matter	"Joint" negative associations
Left superior parietal cortex	Right Pallidum	"Joint" negative associations
Left superior parietal cortex	Right Putamen	"Joint" negative associations
Left superior parietal cortex	Right Thalamus Proper	"Joint" negative associations, Negative associations only
Left superior parietal cortex	Right Ventral Diencephalon	"Joint" negative associations, Negative associations only
Left superior parietal cortex	Right choroid plexus	"Joint" negative associations
Left superior parietal cortex	Right inferior parietal cortex	"Joint" negative associations
Left superior parietal cortex	Right insula white matter	"Joint" negative associations

Left superior temporal cortex	Right bankssts white matter	Positive associations only, "Joint" positive associations
Right pericalcarine cortex	Right frontal pole white matter	"Joint" negative associations
Left inferior temporal white matter	Left parsopercularis white matter	Negative associations only
Left lateral occipital white matter	Right inferior temporal white matter	Positive associations only, "Joint" positive associations
Left posterior cingulate white matter	Central Corpus Callosum	Positive associations only, "Joint" positive associations
Left rostral middle frontal white matter	Brain Stem	Positive associations only, "Joint" positive associations
Left superior frontal white matter	Brain Stem	Positive associations only, "Joint" positive associations
Left transverse temporal white matter	Right rostral middle frontal white matter	Positive associations only, "Joint" positive associations
Right middle temporal white matter	Right frontal pole white matter	"Joint" negative associations

F.3 Obsessions

Table F.3: Selected connections for Obsessions, as well as the model in which they were selected.

Connections		**Models that selected the connections**
Brain Stem	Mid Posterior Corpus Callosum	Positive associations only, "Joint" positive associations
Brain Stem	Right insula cortex	"Joint" positive associations
Brain Stem	Right transverse temporal cortex	Positive associations only, "Joint" positive associations
Mid Posterior Corpus Callosum	Right Thalamus Proper	Positive associations only
Posterior Corpus Callosum	Mid Posterior Corpus Callosum	Positive associations only
Left Caudate	Left post central cortex	"Joint" negative associations, Negative associations only
Left Cerebellum Cortex	Central Corpus Callosum	Positive associations only
Left Cerebellum Cortex	Mid Posterior Corpus Callosum	Positive associations only, "Joint" positive associations
Left Putamen	Posterior Corpus Callosum	"Joint" positive associations
Left Thalamus Proper	Central Corpus Callosum	Positive associations only, "Joint" positive associations

Left Thalamus Proper	Mid Posterior Corpus Callosum	"Joint" positive associations
Left Thalamus Proper	Left superior parietal cortex	"Joint" negative associations, Negative associations only
Left Thalamus Proper	Right precuneus cortex	"Joint" negative associations, Negative associations only
Left Ventral Diecenphalon	Central Corpus Callosum	Positive associations only
Left Ventral Diecenphalon	Mid Posterior Corpus Callosum	Positive associations only, "Joint" positive associations
Left Ventral Diecenphalon	Right Ventral Diencephalon	Positive associations only, "Joint" positive associations
Right Cerebellum Cortex	Right bankssts white matter	Positive associations only, "Joint" positive associations
Right Hippocampus	Right entorhinal white matter	"Joint" negative associations, Negative associations only
Right Hippocampus	Right frontal pole white matter	"Joint" negative associations, Negative associations only
Right Ventral Diencephalon	Right caudal anterior cingulate cortex	Positive associations only
Right choroid plexus	Right superior frontal cortex	"Joint" negative associations, Negative associations only
Right choroid plexus	Right rostral middle frontal white matter	"Joint" negative associations, Negative associations only
Right choroid plexus	Right superior frontal white matter	"Joint" negative associations, Negative associations only
Left post central cortex	Right paracentral cortex	"Joint" negative associations, Negative associations only
Left posterior cingulate cortex	Right Ventral Diencephalon	"Joint" positive associations
Left posterior cingulate cortex	Left lateral occipital white matter	Positive associations only, "Joint" positive associations
Left superior parietal cortex	Right Thalamus Proper	"Joint" negative associations, Negative associations only
Left superior parietal cortex	Right Ventral Diencephalon	"Joint" negative associations, Negative associations only
Left superior temporal cortex	Right Caudate	Positive associations only
Left superior temporal cortex	Right bankssts white matter	"Joint" positive associations
Right insula cortex	Right parahippocampal white matter	Positive associations only
Right insula cortex	Right precuneus white matter	"Joint" positive associations

Right transverse temporal cortex	Right superior parietal white matter	"Joint" positive associations
Left inferior parietal white matter	Right insula cortex	"Joint" positive associations
Left rostral middle frontal white matter	Left transverse temporal white matter	Positive associations only, "Joint" positive associations
Left rostral middle frontal white matter	Right bankssts white matter	Positive associations only, "Joint" positive associations
Left superior frontal white matter	Left transverse temporal white matter	Positive associations only, "Joint" positive associations
Left transverse temporal white matter	Brain Stem	Positive associations only
Right bankssts white matter	Right precentral white matter	"Joint" positive associations
Right bankssts white matter	Right superior parietal white matter	Positive associations only

F.4 Compulsions

Table F.4: Selected connections for Obsessions, as well as the model in which they were selected.

Connections		**Models that selected the connections**
Brain Stem	Right precentral white matter	Positive associations only, "Joint" positive associations
Left Accumbens area	Left lingual white matter	"Joint" negative associations
Left Accumbens area	Left temporal pole white matter	"Joint" negative associations
Left Lateral Ventricle	Right medial orbito frontal white matter	"Joint" negative associations
Right Hippocampus	Right entorhinal white matter	"Joint" negative associations, Negative associations only
Left middle temporal cortex	Left parsopercularis white matter	"Joint" negative associations
Left parahippocampal cortex	Left cuneus white matter	Positive associations only, "Joint" positive associations
Left precuneus cortex	Right Pallidum	"Joint" negative associations
Left precuneus cortex	Right Thalamus Proper	"Joint" negative associations
Left precuneus cortex	Right inferior parietal cortex	"Joint" negative associations
Left precuneus cortex	Right lateral orbito frontal white matter	"Joint" negative associations

Left precuneus cortex	Right medial orbito frontal white matter	"Joint" negative associations
Left rostral anterior cingulate cortex	Right Hippocampus	"Joint" negative associations
Left superior parietal cortex	Right Pallidum	"Joint" negative associations
Left superior parietal cortex	Right Putamen	"Joint" negative associations, Negative associations only
Left superior parietal cortex	Right Thalamus Proper	"Joint" negative associations
Left superior parietal cortex	Right Ventral Diencephalon	"Joint" negative associations, Negative associations only
Left superior parietal cortex	Right insula white matter	"Joint" negative associations
Left superior parietal cortex	Right paracentral white matter	"Joint" negative associations
Right parsopercularis cortex	Right paracentral white matter	"Joint" negative associations
Right pericalcarine cortex	Right frontal pole white matter	"Joint" negative associations
Left inferior temporal white matter	Left parsopercularis white matter	Negative associations only
Left lateral occipital white matter	Right inferior temporal white matter	Positive associations only, "Joint" positive associations
Left rostral anterior cingulate white matter	Central Corpus Callosum	Positive associations only, "Joint" positive associations
Left superior frontal white matter	Brain Stem	Positive associations only, "Joint" positive associations
Left transverse temporal white matter	Right rostral middle frontal white matter	Positive associations only, "Joint" positive associations
Right lateral orbito frontal white matter	Right lingual white matter	"Joint" negative associations

F.5 STAI

Table F.5: Selected connections for STAI, as well as the model in which they were selected.

Connections		**Models that selected the connections**
Left Accumbens area	Right superior temporal white matter	Positive associations only, "Joint" positive associations
Left Caudate	Right inferior parietal white matter	Positive associations only, "Joint" positive associations

Left Caudate	Right transverse temporal white matter	Positive associations only, "Joint" positive associations
Left Inferior Lateral Ventricle	Right post central cortex	Positive associations only, "Joint" positive associations
Left Putamen	Left frontal pole cortex	"Joint" negative associations
Left Thalamus Proper	Left cuneus cortex	"Joint" negative associations
Left Thalamus Proper	Right precuneus cortex	"Joint" negative associations, Negative associations only
Left Thalamus Proper	Right superior parietal cortex	"Joint" negative associations, Negative associations only
Right Cerebellum White Matter	Right cuneus cortex	"Joint" negative associations, Negative associations only
Right Cerebellum White Matter	Right middle temporal white matter	Negative associations only
Right Hippocampus	Right entorhinal white matter	"Joint" negative associations, Negative associations only
Right Thalamus Proper	Right cuneus cortex	"Joint" negative associations
Right Thalamus Proper	Right superior parietal cortex	Negative associations only
Left fusiform cortex	left entorhinal white matter	"Joint" negative associations
Left fusiform cortex	Left lateral occipital white matter	Positive associations only, "Joint" positive associations
Left inferior parietal cortex	Right insula white matter	Negative associations only
Left inferior temporal cortex	Left precentral white matter	"Joint" negative associations
Left isthmus cingulate cortex	Left pars triangularis white matter	"Joint" negative associations, Negative associations only
Left lateral occipital cortex	Mid Posterior Corpus Callosum	Positive associations only, "Joint" positive associations
Left lateral occipital cortex	Right Pallidum	Positive associations only, "Joint" positive associations
Left lateral occipital cortex	Right superior temporal cortex	Positive associations only, "Joint" positive associations
Left lateral occipital cortex	Left pericalcarine white matter	"Joint" negative associations, Negative associations only
Left lateral occipital cortex	Right superior temporal white matter	Positive associations only, "Joint" positive associations
Left medial orbito frontal cortex	Right parahippocampal white matter	Positive associations only, "Joint" positive associations
Left middle temporal cortex	Left precentral white matter	"Joint" negative associations
Left parsorbitalis cortex	Left parsorbitalis white matter	"Joint" negative associations

Left parsorbitalis cortex	Right postcentral white matter	Positive associations only, "Joint" positive associations
Left precuneus cortex	Right Thalamus Proper	"Joint" negative associations, Negative associations only
Left precuneus cortex	Right inferior parietal cortex	Negative associations only
Left superior parietal cortex	Right Thalamus Proper	"Joint" negative associations, Negative associations only
Left superior parietal cortex	Right Ventral Diencephalon	"Joint" negative associations, Negative associations only
Left superior parietal cortex	Right inferior parietal cortex	Negative associations only
Left superior parietal cortex	Right insula white matter	"Joint" negative associations
Left temporal pole cortex	Left superior parietal white matter	Positive associations only, "Joint" positive associations
Left transverse temporal cortex	Left postcentral white matter	Positive associations only, "Joint" positive associations
Right entorhinal cortex	Right superior frontal white matter	Positive associations only, "Joint" positive associations
Right inferior parietal cortex	Right entorhinal white matter	Negative associations only
Right inferior temporal cortex	Right frontal pole white matter	Negative associations only
Right lateral orbito frontal cortex	Right entorhinal white matter	Negative associations only
Right pars triangularis cortex	Right pars triangularis white matter	"Joint" negative associations
Right superior frontal cortex	Right precuneus white matter	Positive associations only, "Joint" positive associations
Left inferior temporal white matter	Left parsopercularis white matter	"Joint" negative associations
Left lateral occipital white matter	Left rostral middle frontal white matter	Positive associations only, "Joint" positive associations
Left medial orbito frontal white matter	Right pars triangularis white matter	Positive associations only, "Joint" positive associations
Left postcentral white matter	Right caudal middle frontal cortex	Positive associations only, "Joint" positive associations
Left precuneus white matter	Right Thalamus Proper	"Joint" negative associations, Negative associations only
Left precuneus white matter	Right inferior parietal cortex	Negative associations only
Left precuneus white matter	Right lateral orbito frontal white matter	Negative associations only

Left temporal pole white matter	Right insula white matter	Positive associations only, "Joint" positive associations
Right bankssts white matter	Right superior parietal white matter	Positive associations only, "Joint" positive associations

F.6 MFI

Table F.6: Selected connections for MFI, as well as the model in which they were selected.

Connections		**Models that selected the connections**
Brain Stem	Right insula cortex	Positive associations only, "Together" positive associations
Mid Posterior Corpus Callosum	Right Cerebellum Cortex	Positive associations only, "Together" positive associations
Left Caudate	Right Putamen	Positive associations only, "Together" positive associations
Left Caudate	Left caudal middle frontal white matter	Positive associations only, "Together" positive associations
Left Cerebellum Cortex	Left superior frontal white matter	Positive associations only, "Together" positive associations
Left Lateral Ventricle	Left Cerebellum Cortex	Positive associations only, "Together" positive associations
Left Thalamus Proper	Right precuneus cortex	"Together" negative associations, Negative associations only
Left Thalamus Proper	Right superior parietal cortex	"Together" negative associations, Negative associations only
Right Cerebellum White Matter	Right middle temporal white matter	"Together" negative associations
Right Hippocampus	Right entorhinal white matter	"Together" negative associations
Right Thalamus Proper	Right medial orbito frontal cortex	Positive associations only, "Together" positive associations
Right Thalamus Proper	Right superior parietal cortex	"Together" negative associations
Left inferior parietal cortex	Left paracentral white matter	Positive associations only, "Together" positive associations
Left insula cortex	Left inferior parietal white matter	Positive associations only
Left paracentral cortex	Right paracentral cortex	"Together" negative associations
Left paracentral cortex	Left precuneus white matter	Positive associations only

Left paracentral cortex	Left superior parietal white matter	Positive associations only, "Together" positive associations
Left post central cortex	Right paracentral cortex	"Together" negative associations
Left post central cortex	Right supra marginal cortex	"Together" negative associations
Left posterior cingulate cortex	Left lateral occipital white matter	Positive associations only, "Together" positive associations
Left precentral cortex	Right paracentral cortex	"Together" negative associations
Left precuneus cortex	Right Thalamus Proper	"Together" negative associations, Negative associations only
Left precuneus cortex	Right superior parietal cortex	"Together" negative associations
Left precuneus cortex	Right supra marginal cortex	"Together" negative associations
Left superior parietal cortex	Right Thalamus Proper	"Together" negative associations, Negative associations only
Left superior parietal cortex	Right choroid plexus	"Together" negative associations
Left superior parietal cortex	Right post central cortex	"Together" negative associations
Left superior parietal cortex	Right supra marginal cortex	"Together" negative associations
Left superior parietal cortex	Right insula white matter	"Together" negative associations, Negative associations only
Left superior temporal cortex	Left inferior temporal white matter	Positive associations only
Left superior temporal cortex	Left postcentral white matter	Positive associations only
Right isthmus cingulate cortex	Right supra marginal cortex	"Together" negative associations
Right precuneus cortex	Right frontal pole cortex	Positive associations only
Right temporal pole cortex	Right superior frontal white matter	Positive associations only, "Together" positive associations
Left fusiform white matter	Left transverse temporal white matter	"Together" negative associations
Left inferior parietal white matter	Right superior temporal cortex	Positive associations only
Left inferior parietal white matter	Left postcentral white matter	Positive associations only, "Together" positive associations
Left precuneus white matter	Right Cerebellum Cortex	Positive associations only, "Together" positive associations
Right inferior parietal white matter	Right supra marginal white matter	Positive associations only
Right middle temporal white matter	Right parsorbitalis white matter	"Together" negative associations, Negative associations only

Appendix G

Assessing the impact of correcting for motion on the results

As mentioned in the section 4.3.3 of the methods, CPM predicts motion from connectivity matrices (see Table 4.2). As a consequence, an additional correction was performed. Partial correlation was therefore used to delete the effects of motion in the prediction of the behavioral measures. This appendix reviews the impact of using partial correlation, and thus how motion affected the prediction of each of the different behaviors.

For BDI, the number of p-values below 0.05 increased, going from 37.5% without partial correlation to 45% with it. The number of p-values below the Bonferroni threshold remained the same, 1.7% .

Regarding OCDS, an increase was also noticed with the application of partial correlation, especially for the p-values below the Bonferroni threshold. Prior to partial correlation, only 10.7% of these p-values fell below the threshold. However, after employing partial correlation, this percentage increased to 23.3%.

For Obsessions, the results obtained did not vary much. In fact, 69.4% of the p-values were below 0.05 before the use of partial correlation, and 71.6% after. The percentage of p-values decreased from 33.9% to 31.7%.

For Compulsions, the percentage of p-values below 0.05 decreased with the use of partial correlation, going from 57.1% to 50%. However, the percentage of p-values below the Bonferroni threshold increased from 7.1% to 13.3%.

For STAI, both the percentage of p-values below 0.05 and below the Bonferroni threshold increased with the use of partial correlation. It went respectively from 44.6% to 48.2%, and from 7.1% to 13.3%.

Finally, regarding MFI,the percentages of p-values below 0.05 decreased and with the percentage below the Bonferroni threshold increased. 66.1% of p-value were below 0.05 before the use of partial correlation, and 7.1% were below the Bonferroni threshold. After employing partial correlation, 56.7% of the p-values were below 0.05, and 15% below the Bonferroni threshold..

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