Machine Learning applications for characterizing brain-damaged patients' level of consciousness

Georgios Antonopoulos

Under the direction of Prof. Steven Laureys and Prof. Christophe Phillips

Coma Science Group
GIGA Consciousness
University of Liège

Thèse présentée en vue de l'obtention du grade de
Docteur en Sciences biomédicales et pharmaceutiques

Liège, 2019

Supervisors

Prof. Steven LAUREYS

Prof. Christophe PHILLIPS

Assessment committee

Dr. Athena DEMERTZI (secretary)

Prof. Roland HUSTINX (president)

Dr. Mohamed Ali BAHRI

Dr. Quentin NOIRHOMME

Dr. Emmanuel STAMATAKIS (University of Cambridge)

Prof. Danilo BZDOK (RWTH Aachen University)

page1

Contents

List of Figures

List of Tables

Acronyms

AUC Area Under the ROC Curve 37, 63, 66

BOLD Blood Oxygen-Level Dependent 13, 45

CRS-R Coma Recovery Scale-Revised 7, 45, 47, 48, 54, 61

CS Classical Statistics 26, 27

CSF Cerebrospinal Fluid 45

CT Computed Tomography 15

CV Cross Validation 35, 36, 38

DMN Default Mode Network 18, 43, 49

DOC Disorders of Consciousness 5, 7, 8, 18, 19, 21, 25–27, 38–40, 43, 44, 51, 52, 59, 60, 70, 75, 76, 78

DT Decision Trees 31, 32, 34, 35

EEG Electroencephalography 27, 44

EMCS Emerged Minimally Conscious State 6, 7, 20, 44, 50

FDG 18F fluorodeoxyglucose 15, 19, 38, 45, 59, 62, 70

fMRI Functional MRI 11, 13, 17–19, 21, 38, 39, 43–45, 75, 77

FN False Negative 36, 72

FN Functional Neuroimaging 8, 12, 17–19, 21, 26, 27, 40, 60

FP False Positive 36, 72

GLM General Linear Model 26, 27

ICA Independent Component Analysis 46

LIS Locked-In Syndrome 6, 7

MCS Minimally Conscious State 6–8, 20, 36, 43–45, 47–52, 54, 59–61, 63–67, 69, 70, 72, 75–77

ML Machine Learning 26, 27, 38, 40

MNI Montreal Neurological Institute 16, 45, 62, 70

MRI Magnetic Resonance Imaging 12, 16, 44, 76

PET Positron Emission Tomography 11, 12, 14, 15, 18, 19, 21, 38, 39, 44, 45, 54, 59–62, 75, 76, 78

RFC Random Forest Classifier 31, 34, 35

ROC Receiver Operating Characteristic 37

ROI Region Of Interest 47

rsfMRI Resting State fMRI 17, 44, 45

RUV Relevance Uptake Value 78

sMRI Structural MRI 13

SPM Statistical Parametric Mapping 26, 45, 60, 62, 65, 70

SUV Standardized Uptake Value 15, 16, 59, 60, 62, 65, 70, 76–78

SVM Support Vector Machine 28, 30, 43, 47–49, 59, 63, 64, 66–69, 72, 77

TMS Trancranial Magnetic Stimulation 44

TN True Negative 36

TP True Positive 36, 49

TPM Tissue Probability Maps 62

VS/UWS Vegetative State/Unresponsive Wakefulness Syndrome 5–8, 18–20, 36, 43–45, 47–49, 51, 52, 54, 60, 61, 63, 64, 66–70, 72, 73, 75–77

Xtrees Extremely randomized trees 31, 34, 35, 63, 64, 66, 68–73, 77

Abstract

Consciousness is the result of an extremely complicated brain function. The exact functionality of the brain resulting in consciousness remains unsolved. Combined forces from many different scientific fields are working on this to get a better understanding on consciousness and its disorders. Medicine, neuropsychology, mathematics and biology are only a few of those fields. Specifically, the medical model can provide us with unique insights as to the functions of typical states of consciousness.

This thesis is focusing on patients with disorders of consciousness. This kind of patients are brain-lesioned individuals which in numerous cases are incapable of responding to requests, despite the fact that they might still have preserved conscious functions. Often, the remaining functionality of a brain is sufficient for perceiving and decoding the surrounding environment or the position of patients in it. Nowadays, we know that lack of responses do not necessarily indicate lack of consciousness. Behavioural-assessment scales for the evaluation of consciousness often provide a vague diagnosis. Mis-diagnosis of consciousness raises clinical as well as ethical issues.

Functional neuroimaging can be used to address this problem by providing an inner overview of the brain functionality of patients with disorders of consciousness. Functional Magnetic Resonance Imaging and Positron Emission Tomography are two commonly used modalities of functional neuroimaging, which are used in the projects of this thesis. They provide a quantification of different brain properties in combination with an accurate spatial representation, which makes them a unique source of information. Machine Learning, being part of the wider Artificial Intelligence field, incorporates algorithms that can efficiently handle high-dimensional data. Such algorithms can unveil patterns of data and undercover interactions of brain regions, using data-driven approaches. Additionally, they provide tools that can ensure success in predicting unseen data. Therefore, they can constitute a necessary and complementary tool to classical statistics for the analysis of Functional Neuroimaging data in Disorders of Consciousness.

The combination of behavioural assessments and functional neuroimaging form an extremely important and unique source of information, for both clinical use and the scientific study of consciousness.

The former is showing the thin line between consciousness and un-consciousness and the latter provides the means to explore it.

This thesis aims at providing tools to assist the behavioral diagnosis of consciousness using Machine Learning in functional neuroimaging data from patients with disorders of consciousness. The studies composing it focused mainly on the two groups that are considered to lie on the border line of responsiveness: i) Minimally Conscious State, and ii) Unresponsive Wakefulness State. Two different modalities, which capture different properties of brain function, have been used. At first we used functional Magnetic Resonance Imaging, from which we extracted brain connectivity features. To those features we applied machine learning techniques to identify the contribution of brain networks to the classification of patients. In the second project, we used the metabolic activity of the brain extracted from Positron Emission Tomography, to classify patients with brain lesions and extract regional information. We applied certain practices, in order to overcome problems such us noisy images, redundant features and limited samples.

Both projects are highlighting these brain regions with the maximum contribution to the classification process, assuming that they are significant to higher order cognitive functions, therefore shedding light on the mechanistic counterpart of the phenomenon of consciousness.

Chapters' Preview

Chapter 1 describes the main terminology, fundamental principles and the disorders of consciousness. The behavioural characteristics of each state of consciousness are presented.

Chapter 2 is based on book chapter **Measuring Consciousness Through Imaging** [1]. It provides an overview of Functional Neuroimaging and how it has been used so far in the investigation of Disorders of Consciousness. The operational principles of the functional MRI and PET are briefly described. The necessary preprocessing steps, which prepare the scans for further analysis are also reported. Finally, some important findings of Functional Neuroimaging, which are used in the later chapters are provided.

Chapter 3 is an overview of the analysis that has been used for data of Disorders of Consciousness (DOC) patients, based in the review work **"Look at my classifier's result": Disentangling unresponsive from (minimally) conscious patients** [2]. Additionally, the data analysis tools that were used in the conducted research are included in this chapter. Statistical methods referred to in the ensuing chapters are only briefly referenced due to the fact that they are well known, validated and considered detrimental today. Algorithms and techniques from Machine Learning, which are applied in next chapters of this thesis are described in more details.

Chapter 4 is based on **Intrinsic functional connectivity differentiates minimally conscious from unresponsive patients** [3], which describes the classification process of Disorders of Consciousness patients using Resting State fMRI (rsfMRI). It starts with the preprocessing followed by the steps we performed to build the model for classification. Extraction and evaluation of features, evaluation and validation of the produced model and its evaluation with a generalized unseen population is described in this chapter.

In Chapter 5, the analysis of PET using Machine Learning is presented. It consists of the preliminary results of the classification of lower states of Disorders of Consciousness.

This work is an effort to create a diagnostic tool with minimum human intervention but also to study the neural correlates of consciousness. The process of estimation of models' parameters and the models' validation with the unseen data are presented.

The last chapter contains the conclusion, drawbacks and general considerations drawn from the research conducted in the chapters 4 and 5.

Chapter 1

Disorders of consciousness and means of evaluation

Summary

Although "consciousness" is lacking a concrete definition, for the sake of scientific progress, scientists have come up with ways to define and quantify it. A self-explanatory start for the research of consciousness is to observe patients for their behavioral symptoms in order to accurately assess in the different states of pathological (un)consciousness. Then we can study the pathogenesis of it. Behavioral assessments aim at assigning the behavioral characteristics to a level of consciousness. It is the main way to evaluate if a patient with brain lesion has consciousness and in what degree.

Keywords*: Consciousness, Disorders of Consciousness, Behavioural assessments, CRS-R.*

1.1 Consciousness

There is a big paradox in the scientific studies of consciousness: we are investigating something for which there is no global or commonly accepted definition. If we would let this inconvenience hold back our passion to investigate this extremely fascinating topic, we would not have started at all and the brain research would not have had the very important findings that came up while "looking for consciousness". Additionally, it is also possible that a global definition will come up as a result of research findings.

Following the neurological definition of consciousness framed by Plum and Posner[4], Laureys in 2005 [5] suggested that consciousness has two components as illustrated in figure 1.1.
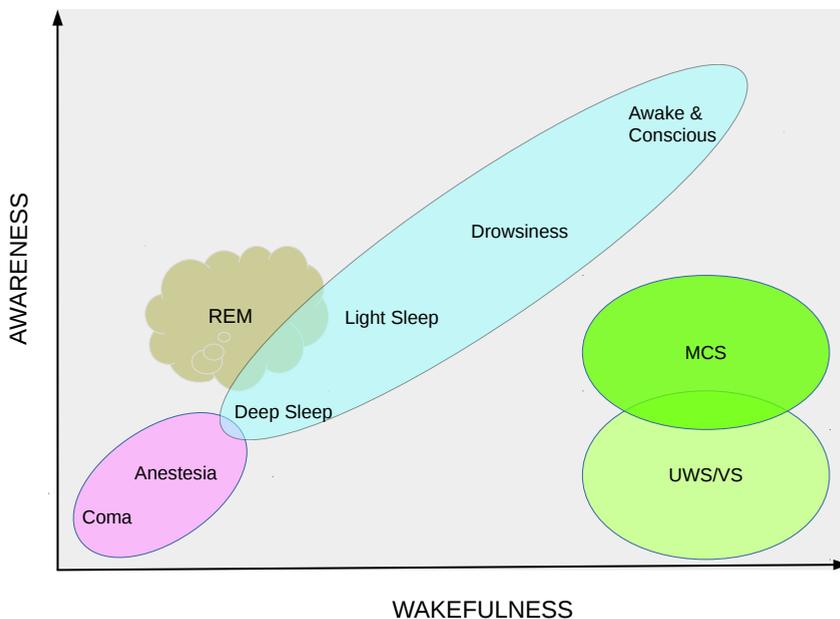
Figure 1.1: The two components of consciousness. Consciousness as a result of Awareness and Wakefulness. Projecting consciousness in two components permits the necessary quantification and definition for researchers to further explore consciousness. Adjusted with permission from [5]

Wakefulness describes arousal, vigilance or simply eye opening. Awareness, that can be divided in internal awareness, which is associated to internal thoughts and processes and external awareness

that is associated to external inputs or stimulus of the environment [6].

To better illustrate the idea of the two-components model for describing consciousness one can think of sleep, excluding the REM phase. In a healthy consciousness state, the more we move to the right of the plot, thus increasing awareness, the more "awake" we become. One exeption to this is REM sleep, in which we dream (and can have recollections aka awareness of it) while we are asleep. Reversely, decreasing from wakefulness to drowsiness, light sleep, deep sleep, sedation/anesthesia means immediately a drop in the level of awareness. Here, we focus on the conditions that fall outside the "main line" of the plot for non pharmacological or physiological reasons, they are called DOC and describe low awareness in full wakefulness. DOC are resulting from brain lesions caused by internal (e.g. anoxia) or external (e.g. traumatic brain injury) factors.

In other words, being more "awake" is not followed by an increase of the perception of personal status or surrounding environment. When this phenomenon occurs, we are talking about DOC. Differently, DOC are referring to a deficient level of consciousness that follows a period of coma. The following paragraphs present the different pathological states of consciousness.

1.2 Defining disorders of consciousness

Coma is the state of a person who is completely unresponsive and cannot be awakened. The main causes of coma are brain trauma, lack of oxygen (anoxia), brain stroke or intoxication for example with drugs. During coma, patients cannot respond to any stimulation and will only show reflexes. Usually coma lasts from some hours up to several weeks.

After this period, coma patients might die or wake up showing normal sleep cycles with eyes opening. However, although awake, they might not necessarily recover any non-reflexive behavior or communication. They might immediately wake up in any of the different states of consciousness or slowly progress to it.

Vegetative State/Unresponsive Wakefulness Syndrome (VS/UWS) [7] is observed when patients recover sleep cyrcles while they show only reflexive behavior. It is characterized by the absence of any kind of command following, lack of self or environmental awareness and it is believed to have incomplete perception of pain [8]. If VS/UWS has a traumatic aetiology and the patient remains in this condition for more than one year, or in case it is non traumatic and the patient remains VS/UWS for more than three months then it is called Permanent Vegetative State [9, 10].

Minimally Conscious State (MCS) is characterized by signs of consciousness of self and/or environment. Patients are not able to communicate functionally but they are able to show language-independent signs of consciousness like visual pursuit or provide responses to simple commands either verbally or in other ways. Reproducibility of responses is mandatory in order to ensure intentionality [11]. Non-reflexive reactions, such as mirror or object tracking, and emotional behaviors, like crying or laughing, are often observed in MCS patients. MCS, just like VS/UWS can be permanent or a transition to another state of consciousness.

Once patients are able to use objects in the appropriate way or are able to communicate they are characterized as being in **Emerged Minimally Conscious State (EMCS)** [11].

Locked-In Syndrome (LIS), though is not a disorder of consciousness, it is worth mentioning due to the fact that it can often have similar behavioral signs and thus clinical evaluation results as VS/UWS [12]. LIS patients are characterized by the presence of eyes opening, aphonia, tetraplegia or quadriparesis, preserved awareness of self and environment and the ability to communicate via eye blinking or eye movement as described in the American Congress of Rehabilitation Medicine in 1995. The insufficiency of behavioral tools to diagnose LIS necessitates the development of paramedical tools.

1.3   Detecting consciousness through behavior

The precise diagnosis of the state of consciousness is of great importance for medical, ethical and scientific reasons. With an accurate diagnosis the appropriate medication and medical care will be provided from caregivers to patients. For example, upon painful stimulation, MCS patients activate the pain matrix in a similar way as healthy subjects, while in VS/UWS patients the activation stops before the higher-order integration cortices. Therefore we suspect that VS/UWS patients do not integrate pain in an efficient way, and therefore are not aware of it. However, if VS/UWS patients show signs of discomfort, pain treatment is provided [13, 14]. This results in the improvement of the quality of life of patients. Another ethical aspect of the accurate detection of consciousness is the end-of-life issue as the presence or lack of consciousness is a factor of great importance in such a debate. The importance of a correct diagnosis is also increased by the fact that there is a different prognosis between MCS and VS/UWS patients [15]. Patients in MCS have better chances to recover consciousness compared to those in VS/UWS. An accurate diagnosis would also help scientists to produce more reliable results and thus create more reliable treatment and rehabilitation methods.

In centers that are not specialized in brain traumas and deficits of consciousness, diagnosis of DOC patients is based on doctors opinion just by observing the patients behaviorally, although several standardized behavioral scales have been proposed and tested. Some of the most widely used scales for assessing the level of consciousness are the Coma Recovery Scale-Revised (CRS-R) [16], the Full Outline of UnResponsiveness scale (FOUR) [17] and the Coma Glasgow Scale (CSG)[18]. To date the most sensitive scale for assessing DOC and especially disentangling VS/UWS from MCS [19] is CRS-R. The CRS-R has excellent content validity and is the only scale fulfilling all of the Aspen Workgroup criteria for good standardized administration and scoring [11]. As all patients' level of consciousness, from whom the data has been used for this thesis, have been evaluated using the CRS-R, following paragraphs will give a brief overview of it.

CRS-R contains 23 items, that are used to assess auditory, visual, motor, oromotor, communication and arousal functions. Also, brain stem reflexes are assessed in order to assist in the interpretation of some subscale items and to evaluate brain stem lesions and help to detect LIS patients. Repetitive assessments ensure that responses are consistent reflexes and are indeed the result of cognitive function. The items in the subscales are hierarchically organized with the lowest items being associated with reflexes and the highest ones with cognitively-mediated behaviors. The auditory subscale includes four items that range from response to auditory stimulation in the lower level to understanding of simple language. The visual subscale includes five items starting from visual startle to object localization and recognition. The six-items motor function subscale is the one with the most items. It includes in the lowest level abnormal posturing and in the highest, that denotes EMCS, the functional use of objects. The oromotor subscale contains three items: oral reflexive movement, vocalization and intelligible verbalization. The communication subscale describes whether a patient can have intentional/functional communication. Patients are questioned about personal and situational matters. If communication is achieved the patient is diagnosed directly as EMCS but if not, the diagnosis is based on the results of the other subscales. Finally, the arousal subscale assesses alertness during the examination. CRS-R is the most sensitive scale for a differential diagnosis between VS/UWS, MCS and EMCS. It is essential for evaluating the state of consciousness, though it is not always successful. Preserved consciousness can be underestimated in up to 40% of patients without the use of standardized behavioral assessments [20] and much less but still significant with the use of them [21–23]. There are some pitfalls that might lead to a false negative misdiagnosis, while false positives (VS/UWS as MCS) are unlikely due to the fact that a patient is characterized with the highest state of consciousness. To

illustrate better, when an item which is denoting an MCS state, such as "consistent movement to command", is scored, then even if all the other items of the other subscales are in line with a VS/UWS, diagnosis of the patient will be MCS. Deafness or blindness of the patient or other cortical sensory impairments, cognitive deficits such as aphasia and apraxia and low vigilance during the assessments might be reasons that non-reflex reactions or voluntary communication are not produced by the patient although preserved consciousness might still exist.

1.4   Summary

The level of consciousness can be approximately quantified and measured using behavioural scales for clinical and scientific purposes. Patients with DOC may have difficulties or being completely unable to perceive any external stimulation, but even if they can, it is possible that they are unable to express it. Thus, behavioural scales cannot always provide an accurate evaluation of consciousness. This complicates the diagnosis for patients with DOC and generates the need for other means to be developed, such as Functional Neuroimaging (FN), to assist the correct evaluation.

Chapter 2

Functional Neuroimaging in Consciousness

Summary

Functional neuroimaging provides a visualization of brain operation mode. Positron Emission Tomography (PET) and Functional MRI (fMRI) can reveal pathways of function or metabolism of different elements in the brain. The representation of this activity in three dimensions allows doctors to perform visual evaluation, but also provide important data for analysis using mathematical tools in computers. Functional neuroimaging has contributed in mapping cognitive tasks to brain networks and regions, and thus it provided new aspects in research. The operational principles of each modality, an overview of data preparation which precedes the main analysis and main brain networks are presented in this chapter.

Keywords*: Functional Neuroimaging, MRI-fMRI, PET, image preprocessing, brain networks.*

2.1   Principles of Functional Neuroimaging

FN techniques provide important information about the operation mode and functional properties of the brain. In vivo brain imaging techniques can be used to evaluate how active brain regions are, how they connect and how they interact with each other. Magnetic Resonance Imaging (MRI) and PET are two of the most commonly used imaging techniques for the investigation of consciousness and the disorders of it.

2.1.1   (Functional) Magnetic Resonance Imaging

Originally MRI was called Nuclear MRI but was later renamed to MRI to avoid negative connotations as it is not using any radioactive tracers or X-rays and thus there is no damaging radiation for humans. MRI bases its function on the magnetic properties of atomic nuclei in order to give an insight into human organs and soft tissues.

The key concept, called nuclear induction or nuclear magnetic resonance, is that spinning nuclei of bulk matter absorb and transmit radio waves when being exposed into an external magnetic field (here like in most of the bibliographical reports will be referred as B0) [24]. This idea is applied to hydrogen atoms of the human body due the fact that human tissues contain them in big percentage. Water molecules contain hydrogen nuclei and each nucleus consists of a constantly spinning single proton which creates a magnetic field around it and thus behaves like a small magnet. When the magnetic field B0 is applied protons will align or anti-align with it (figure 2.1).

The amount of the aligned and the anti-aligned protons is almost the same. Once the atoms are spinning in parallel or anti-parallel way to the field, radiofrequency (RF) pulses are applied to B0, inducing a transverse magnetization. When RF pulses have the same frequency as the precessional frequency of the protons, then they: 1) move out of the alignment with B0 and thus to a higher energy level and also 2) synchronize movements with each other. During this process, protons oscillate and create a small transverse magnetic field resulting in the induction of a small current in a receiver coil of the MRI scanner. With the stop of RF pulses protons start to return to the low-energy state by releasing energy to the surrounding matter. The time of longitudinal relaxation, that is actually the time of restoration of longitudinal magnetization, varies between different tissues and is described by the T1 relaxation constant. The time that protons need to fall out of phase, while transverse magnetization is getting lost is described by the transverse or T2 relaxation constant. The transverse relaxation is not caused only by variations of neighboring protons of the tissues but also by inhomogeneities of the
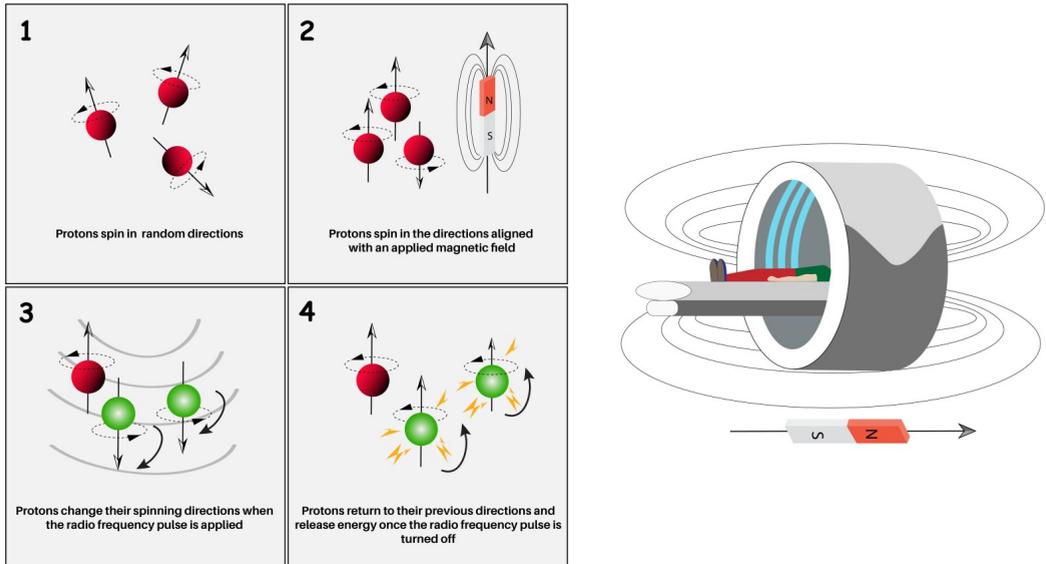
Figure 2.1: Protons magnetization. The different states of protons in an MRI scanner. As hydrogen atoms spin in random directions, they produce a tiny electrical current and axiomatically associate magnetic currents, behaving as tiny bar magnets. In the resting state (1), atoms behave like small magnets with random orientation. (2) When an external magnetic field is applied to the proton, they align or anti-align to the field. (3) When external RF pulses are applied, protons are deflected from previous orientation. (4) When RF pulses stop, protons start to return to their equilibrium state (emit the energy absorbed in previous phase) and have a transverse spinning magnetization. Adapted from `https://knowingneurons.com/2017/09/27/mri-voxels/`

local magnetic field. The relaxation that describes the T2 relaxation in combination with the relaxation caused by the magnetic field inhomogeneities is called T2* (T2star) relaxation.

Structural MRI (sMRI) images make use of differences in protons' relaxation times to contrast different types of brain tissues. A sMRI is a 3D volume that represents the structure of the brain where tissues in this volume are represented by different intensity values. Each 3D volume consists of smaller rectangular cuboid elements called voxels (from volume elements), with edges near 1mm for current research scans (figure 2.2)).

fMRI images are contrasting the T2* relaxation time and are making use of the differences in the magnetic susceptibility between oxygenated and de-oxygenated haemoglobin. An fMRI session consists in the sequential acqusiition of a series of fMR images, where each volume represents a combination of blood concentration and cerebral blood flow (CBF), and oxygen metabolism (CMRO2) called Blood Oxygen-Level Dependent (BOLD) signal. BOLD was first described by Ogawa in 1990 [25]. BOLD contrast images are based on the idea that regional brain activation has increased demands in
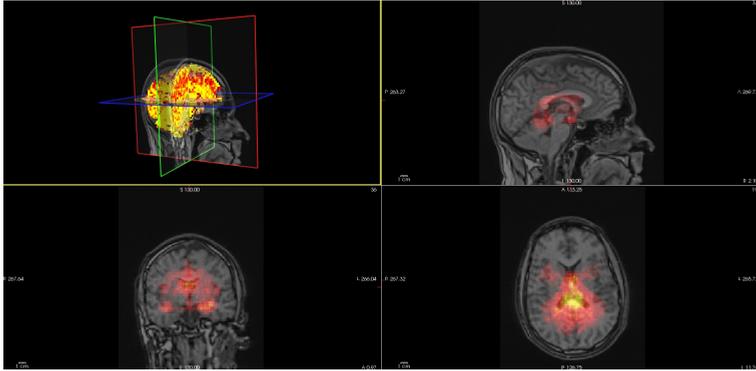
Figure 2.2: Structural and functional MRI volume of healthy control. The structural in gray-scale provides the contrast between tissues, and here is overlapped by one slice of functional MRI, represented in "heat colormap". Top left quarter provides a 3 dimensional view of the scan. The other three represent the sections of the sagittal, the axial and the coronal plane.

sugar and oxygen, which are delivered with blood perfusion, and thus is associated to increased blood flow. By capturing the previous activities of the brain during specific tasks, areas having increased metabolic demands and thus bigger involvement in the task can be pointed out.

2.1.2 Positron Emission Tomography

PET is a nuclear medicine functional imaging technique. Images obtained by a PET scanner can depict blood flow or a biochemical reaction taking part in human or animal bodies. PET measures the spatial distribution of radioactive materials or else radiotracers in organs of live beings. A radiotracer is synthesized by binding a radioactive atom (isotope) to a molecule. When injected in the body, radionuclides in PET-radiotracers emit positrons that annihilates with the surrounding electrons, resulting in the emission of two 511 KeV gamma ray photons. The two gamma ray photons travel in opposite directions and can be detected by the scanner within a finite time interval, usually in the order of 4-12 ns. Image reconstruction algorithms are used to translate the detected couples of photons into 3D volumes (figure 2.3).

The photons captured by the detectors, are not always coming from the ideal aforementioned scenario, but can be the result of the coincidence effect or the scatter effect [26]. The coincidence effect occurs when two unrelated photons are registered as coming from the same event and as a result increases the background signal of the final image. The scatter effect or Compton scattering, is the result of the interaction between photons and charged particles that end up to diverted directions, and

thus photons reaching mis-positioned detectors. The scatter effect has an impact on the overall contrast of the scan. For both events corrections can be applied. Besides errors on the photon detection, PET scans need to be corrected for the photon attenuation. Photon attenuation refers to the effect that photons that are absorbed by the tissues of the body. The attenuation correction is performed using anatomical images, such as Computed Tomography (CT) scans.
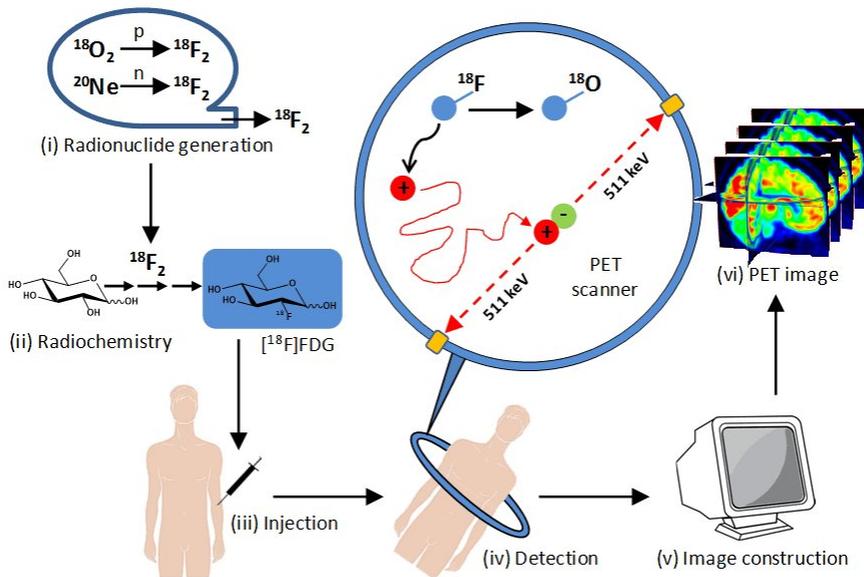
Figure 2.3: Overview of the steps followed for a PET scan. Starting from upper left corner it shows the generation of radionuclides and the radiotracers. The radiotracer is injected in the human subject. When the subject is entered in the scanner radioactivity is captured by the sensors. The final product is a sequence of 3D volumes. Adjusted from `https://www.rah.sa.gov.au/nucmed/PET/pet_docguide.htm`

One of the most common radiotracers in neuroimaging is 18F fluorodeoxyglucose (FDG), which is used to map glucose uptake in the gray matter and can be used to infer the neuronal activity of the brain. In FDG-PET it is a common practice to acquire scans at a specific time after injection, when the concentration of the tracer appears to stabilize temporarily. In this time a few scans are acquired and then averaged to account for subjects' moving noise.

Standardized Uptake Value

The Standardized Uptake Value (SUV) is a method proposed by the Quantitative Imaging Biomarkers Alliance [27]. It rescales the intensities of PET scans, so that they could be used in quantified analysis. It aims at minimizing the variance between scans by incorporating those factors that affect the intensity

of the scans. Three main factors are affecting the concentration of glucose in the brain and thus the captured activity which is reflected to the voxels intensities. One is the time between tracer injection and scanning, which is crucial because of the half-life time of the radioactive materials. The other two parameters are the mass of the subject and the injected dose. The ratio of these two provides a rate of the diffused tracer per body mass units and together with the decay-corrected acquired activity, they provide the SUV. The following formula encompasses the previous:

$$ SUV = \frac{C_t}{\frac{Dose}{BodyMass}}, $$

where C_t is the decay corrected voxel intensity.

2.1.3   Standards in image analysis

Neuroimaging analysis is only possible thanks to the progresses of image and signal processing tools. Images as produced by a scanner, sometimes referred as raw images, need to go through a preparation process that is usually called preprocessing. This prepares them for comparisons with scans of other subjects (inter-subject analysis) or with scans from the same subject (intra-subject analysis), but from different time points or brain states. Preprocessing parameters are not the same for all modalities or pathologies and need to be examined carefully before being applied.

When analyzing scan time-series one should consider and correct for movements of a patient, a process known as motion correction or realignment. Subjects cannot be placed in the scanner in the exact same position and also have different physical characteristics like weight, height, size of neck or head. These create differences in the brain orientation of a scan. In order to be able to compare volumes, brain scans need to be in line so that homologous regions spatially overlap. In general, it is common to orient the brain scans manually as a first step, so that they all have the same direction. This process is called reorientation and precedes spatial transformations. Next step is the spatial normalization, aiming at a more accurate overlap of brain regions by morphing volumes in order to minimize structural differences between subjects. In MRI, usually spatial processing is performed in structural images and then applied in the functional ones. Thus, functional and structural scans need to be coregistered. Scans undergo image transformation algorithms to match a brain template, with the Montreal Neurological Institute (MNI) template being the most commonly used. Once images are formed in a way that all brain regions of different scans are overlapping, voxels can be labeled using tissue probability maps according to their structural properties. Some algorithms can perform spatial normalization and tissue segmentation at the same step [28]. Smoothing is usually the last part of

preprocessing and aims to increase the Signal to Noise Ratio (SNR). A typical fMRI preprocessing is illustrated in figure 2.4.
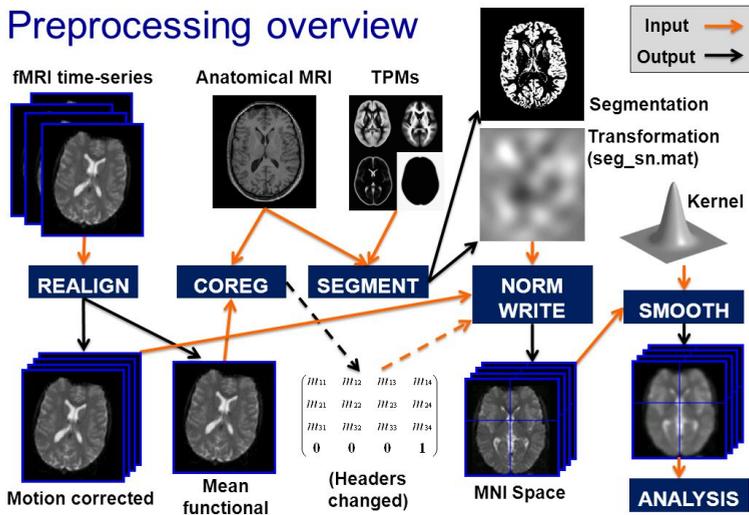
Figure 2.4: An overview of a typical preprocessing pipeline for functional MRI images. Functional MRI volumes are corrected for motions of the subject that occur during data acquisition. Structural MRI is used to estimate spatial transformation and tissue segmentation. The estimated transformation is then applied to functional images after they are coregistered to the structural MRI. Smoothing is following to increase signal to noise ratio. Adapted from SPM course Ashburner, Ridgway (`https://www.fil.ion.ucl.ac.uk/spm/course/slides11/02_Preprocessing_FIL2011May.pptx`)

2.2 Brain networks

The use of FN techniques provided important findings in brain research such as the organization of brain into networks. A "network" is formed by those brain regions which demonstrate a synchronized activity over time, when subjects perform a specific cognitive task. Networks can also be identified in different study designs, such as during resting state, that is the one performed in the data we analyzed for this work. The resting state paradigm refers to task- and stimulus-free data acquisition aiming at capturing a baseline of brain activation [29]. During rsfMRI subjects are instructed only to relax and refrain from sleeping. Other paradigms are "active paradigms" and "passive paradigms". The former refers to the study design where subjects get stimulated during acquisition and aims at capturing the contrast between different conditions of brain activity. In the latter subjects are not instructed to perform any task but they receive different kinds of stimulation, such as music, familiar voices, pain and others.

When it comes to resting state, two well known networks have been suggested to characterize the baseline of brain's activity in rest; the "intrinsic" network, encompassing medial brain areas, and the "extrinsic"network, encompassing lateral frontoparietal areas. These networks appear to be anti-correlated, meaning that when one of them is active the other one "goes down". "Extrinsic" has been associated to environmental perceptions and brain operations associated sensory input such as auditory [30], visual [31]. The "intrinsic" functional network, which coincides with the *Default Mode Network (DMN)*, appears to be the dominant network during rest and has been noticed both using PET [32] and fMRI [33]. Activity in the "intrinsic" network has been associated to internal awareness such us self-related thoughts [6]. It encompasses precuneus/posterior cingulate cortex, mesiofrontal/anterior cingulate cortex and temporoparietal junction areas 2.5. *Salience* network, like DMN, has been associated to "higher order" cognitive functions and encompasses the bilateral dorsal anterior cingulate cortex [34], the frontoinsular cortex, and the frontopolar cortex. It is involved in salient emotional stimuli [35], information perception, response selection [36], pain related processes [37, 38]. *Frontoparietal* network is critical for executive control and decision making in goal-driven tasks [39, 40]. It involves regions in the inferior parietal lobe, dorsal premotor cortex and interparietal sulcus. The *Auditory* network encompasses primary and secondary auditory cortices (including Heschl's gyrus, bilateral superior temporal gyri) and posterior insular cortex. It has significant involvement in audition and sound perception and discrimination [41, 42]. The *Sensorimotor* network is associated to motor tasks [29] and involves somatosensory/midcingulate cortex, motor and middle frontal gyri [43]. The *Visual* network appears in the lateral and medial posterior occipital corices and is linked to simple or complex visual activities but also to Braille reading [41]. It can be divided further in three networks: the lateral visual network, the medial visual network and occipital visual network [43]. All networks are illustrated in figure 2.5.

2.3 Functional neuroimaging in DOC

FN is used to investigate the brain functionality of all consciousness states in order to uncover neuronal correlates of consciousness and to complement the behavioral examination and thus minimize misdiagnosis.

PET studies have been used to investigate brain activity in terms of glucose metabolic activity in DOC patients [44, 45]. Already in 1987, the global glucose metabolism of VS/UWS patients was found to be less than half compared to that of healthy subjects [46, 47]. Acute VS/UWS have been found to have significantly higher metabolic activity in all cortical regions exept frontal lobe, compared
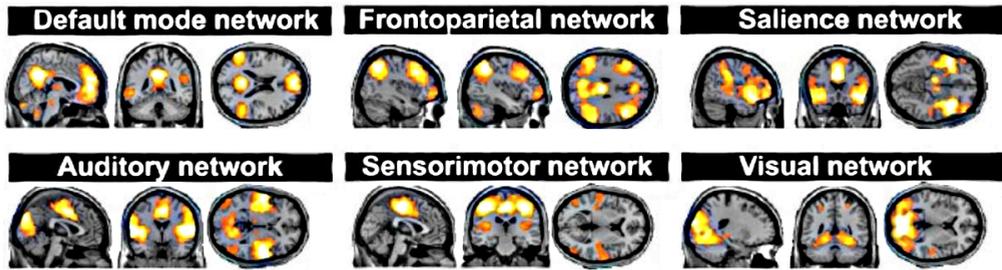
Figure 2.5: Basic brain networks which can be identified using functional neuroimaging in resting state. Here, six networks which reflect "higher" order cognitive function and sensory functions, as they result from the statistical analysis in healthy controls. For visualization purposes they are rendered on structural MRI in Montreal Neurological Institute coordinates. Adjusted from [3]

to permanent VS/UWS [47]. Coma patients and VS/UWS have found to have an overlap in regional metabolic rates [48].

Besides local cerebral activation, PET can also provide information about connectivity between cortical regions as well as connectivity between thalamus and cortex [5, 49–51]. Typical FDG-PET images of healthy subject and DOC patients can be seen in figure 2.6

In fMRI several study designs can be used to investigate subjects brain function in voluntary modulation, spontaneous reactions or in rest. The "active paradigm" design has been used in DOC patients for detecting voluntary brain activation. fMRIm scans acquired during the performance of two different mind wandering tasks were examined by Owen in 2006 [52] in a DOC patient. The patient during scanning was instructed to either imagine playing tennis or walking in the house. Brain activation was very similar to those of healthy subjects concluding that there was command following as measured by willful modulation of brain function, in a patient that was clinically diagnosed as VS/UWS. With some variations of the commands researchers aim to trigger activations in language network [53], in premotor cortex [54] and supplementary motor area [30, 55]. Some studies also tried "passive" paradigm in DOC patients [13, 43, 56] aimming at detecting higher order cortical activation patterns. These activation patterns, which are atypical for unresponsive subjects, could indicate preserved consciousness and constitute or surrogate markers of good prognosis.

Many DOC patients are not able to perceive any commands or stimulation restricting the use of the two aforementioned paradigms. Resting state paradigm can be used to overcome those limitations. FN studies comparing DOC and healthy subjects show significant impairment of the fronto-parietal network encompassing anterior cingulatemesiofrontal and posterior cingulateprecuneus and prefrontal
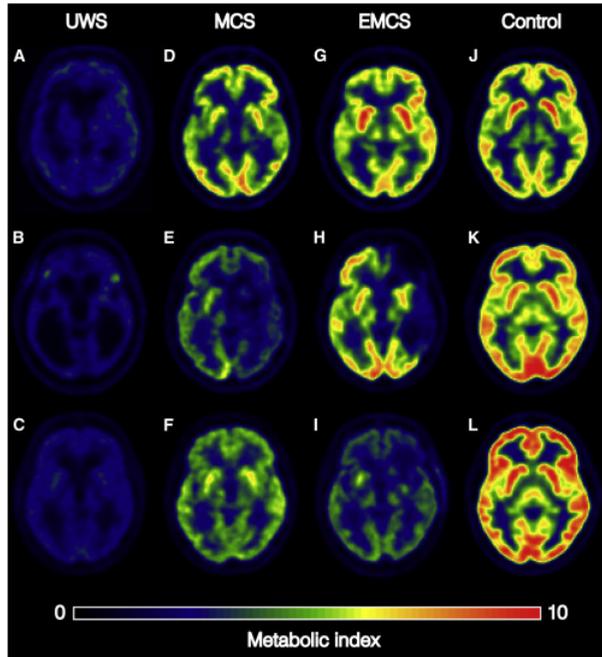
Figure 2.6: FDG-PET scans in different states of consciousness. A-C demonstrate patients ended up in VS/UWS following subarachnoidal hemorrhage (A), traumatic brain injury (B), and anoxic brain injury. In D-F are shown MCS patients following traumatic brain injury (D), hemorrhagic stroke (E), and anoxic brain injury (F). G-I shows EMCS patients following traumatic brain injury with intracerebral hemorrhage (G), traumatic brain injury (H), and anoxic brain injury (I). In J,K and L are shown healthy subjects. Adapted from [50].

and posterior parietal associative cortices [44, 49, 57–60]. These findings were confirmed by both PET and fMRI [61–64] and more than two distinct networks were associated to internal and external awareness [6].

The functional connectivity studies indicate that lack of consciousness is related to a cortico-cortical ([57, 65, 66]) and thalamo- cortical connectivity impairment [65] supporting the hypothesis that consciousness is strongly associated to frontoparietal connectivity [5, 58, 67]. FN studies on conscious perception in healthy volunteers [68, 69] but also in pharmacological loss of consciousness (see [70] for a review) and sleep (see [71] for a review) are in line with the aforementioned theory.

2.4 Summary

FN gives important information on patients' diagnosis as well as for research purposes. It provides an accurate spatial representation of brain functionality in combination with the functional properties. Brain activity as captured with FN has led to the identification of brain operation mode into networks and has opened new directions within brain research. In DOC, the findings complement behavioral assessments of consciousness, and together they form a very important piece of information. This information can be used to to enhance our understanding of the way the brain is functioning but also to improve the accuracy of diagnosis of patients.

Chapter 3

Data analysis methods. From interpretations to predictions

Summary

The majority of knowledge we have so far about the neural correlates of consciousness is based on the analysis with classical statistical methods. Regional activations as portrayed by functional neuroimaging can provide evidence to infer or predict. Classical statistics and Machine Learning methods are used in neurosciences for data analysis. Machine Learning is getting more and more attention in recent years. Such sophisticated methods with complex algorithms incorporate tools which involve data interaction for making a decision. With the time passing, problems like high dimensionality and high computational demand are being solved and small datasets get larger. Therefore, a lot of chances appear for Machine Learning to be applied and thus expand our knowledge in the field of DOC but also to develop tools that help in patients' diagnosis.

Keywords*: Data analysis, Classical Statistics, Classification, Machine Learning, assisting diagnosis.*

3.1 From Statistics to Machine Learning

There is a big debate on what the differences between Classical Statistics (CS) and Machine Learning (ML) are [72–74]. One widely accepted view, which seems to fit with the analysis as it has been performed in FN of DOC over the years, is that CS has been mostly applied on inference for rejecting a hypothesis while ML techniques provide more accurate and generalized predictions on the basis of new data [75]. It is noteworthy that some methods of CS and ML have a common mathematical basis and therefore, in some cases they provide similar results. Additionally, methods of CS are often embedded in steps of ML algorithms which in fact increases the overlap of the two fields.

In FN, the hypothesis-based explanatory nature of statistics is providing answers about brain functions. By performing tests on selected properties of the data, one can reject null-hypothesis and estimate the degree of certainty of this decision. Two of the most common objectives of using CS in FN aim at the detection of brain regions associated to specific tasks and on finding synchronized brain regions that are also called co-activation regional patterns. The latter is also referred to as functional connectivity [76, 77]. Predictions by means of CS have also been applied without, though, having provided any generally applicable tool [52, 55].

Neuroimaging data of DOC patients were, until the beginning of the current decade, almost exclusively analyzed using CS. Toolboxes, like FSL [78] and Statistical Parametric Mapping (SPM) incorporated several methods of CS and provided neuroscientists free, validated and user friendly tools much before those of ML.

The General Linear Model (GLM), implemented in SPM, is the most widely used model for hypothesis testing and statistical modeling in FN. It incorporates linear Algebra and statistical models such as ANOVA, ANCOVA, MANOVA, MANCOVA, t-test, F-test and linear regression (for more details [79, 80]).

In SPM statistical tests are performed across all voxels (mass univariate analysis). If the analysis is performed on the sequence of volumes of one subject, it is referred to as intra-subject (first-level) analysis. In intra-subject analysis, time-series of voxels are being analyzed and aims at finding similar temporal patterns between regions or at associating perception of external stimulis to brain regions. A different type of analysis is performed on data that are acquired from many subjects, which belong to one or more groups. This analysis aims at finding systematic differences of the formed populations and is referred to as inter-subject (second-level) analysis. In both cases GLM, given a defined model, is looking for the parameters that capture the maximum possible variance from the data and fit the data in the model. The estimated parameters represent the contribution of each regressor to the designed

model. An illustration of the GLMcan be found in figure 3.1.
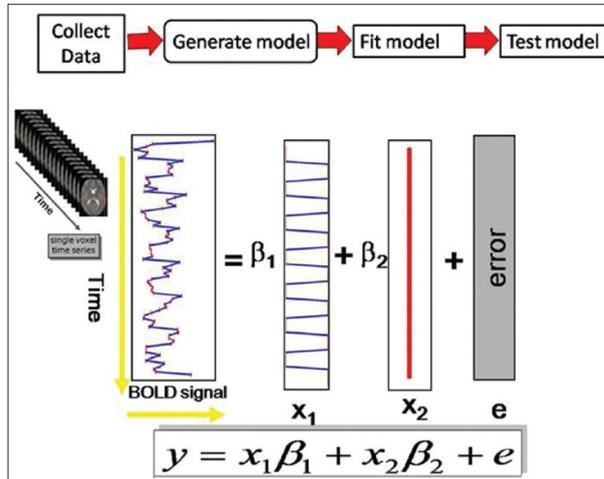
Figure 3.1: General Linear Model applied in fMRI time series. y is the acquired signal. x_i represent the predictors or components that explain the acquired signal. β_i is the contribution weight of each predictor. Finally, *error* is the variance that cannot be explained by the predictors (noise). Adapted from [81]

Though CS methods have provided researchers with important information about neural correlates of consciousness, their diagnostic capacity in terms of predicting unseen data remains poor. It is very common in CS that the whole available dataset is used for inference and thus generalization on unseen data is prohibited. However, in the past, due to the need of having diagnostic tools that complement the behavioral assessments, researchers applied statistics for diagnosing patients at a subject-level [52].

The application of ML techniques is getting more attention in the analysis of DOC mainly because they can provide accurate predictions on "previously unseen data". Applying ML in FN data for classifying DOC patients means simply to look for ways to assign neural correlates to a certain state of consciousness. For both diagnosis and prognosis of DOC patients, ML has been applied to either detect command-following or to differentiate patients using structural or functional properties from data acquired during rest. In command following, the aim is to detect responses from a subject that would indicate an indirect communication which could not be expressed behaviorally. Electroencephalography (EEG) as well as FN modalities have been used for such experiments [30, 52, 55, 82]. In resting-state acquisitions, group comparisons are performed aiming at finding the differences in the baseline brain function of the involved groups of subjects.

3.2   Classifiers

In classification, samples are identified as parts of a specific group of a given population. A classifier is a function that takes categorical or numerical variables as inputs and provides a prediction as an output. Classifiers can be characterized as parametric or non parametric. Parametric classifiers have fewer underlying parameters and can be affected by the violation of normality assumption for the involved populations [83]. Non-parametric classifiers make no assumptions about the data distribution [84]. Classification can be divided in supervised and unsupervised. Supervised classification is the form of learning through examples or labeled data. The process of learning by using labeled data is called training process. In unsupervised classification, also known as clustering, data are not labeled and are classified according to certain similarity indexes [84].

3.2.1   Support Vector Machine

One of the most popular classifiers widely applied in various fields, including Neuroimaging, is the Support Vector Machine (SVM) [85]. A (linear) SVM, given a training set of two groups with D dimensional samples, calculates a hyperplane of $D-1$ dimensions that separates the groups. In the simplified form of two dimensions, the algorithm would look for a line that is splitting the training sets by having the least possible samples on the wrong side. In practice, infinite possible separation hyperplanes could be drawn.

Hyperplane

In an SVM the optimal hyperplane is selected using the idea of maximum-margin, meaning that the selected hyperplane is the one that maximizes the distances between the samples of the two classes (3.2) and the hyperplane. To decide about the hyperplane, the algorithm is choosing some of the samples of the two classes, that are called support-vectors. The word "vector" is used due to the mathematical representation of each sample as a vector. As "support-vectors" are characterized those samples that define the hyperplane. The algorithm maximizes the distance between the hyperplane and the support vectors, when at the same time minimizes a given loss-function [86].

For a more mathematical approach, let $x_i \in \mathbb{R}^p, i = 1, ..., n$ be a training set of n samples and p features and $y_i \in \{-1, 1\}$ the corresponding labels, forming a couple (x_i, y_i) for each sample. The hyperplane is described by the formula: $w^T x + b = 0$, for $w, b \in \mathbb{R}$. Samples falling on one side of the hyperplane are labeled with -1 and on the other side with 1 so that:
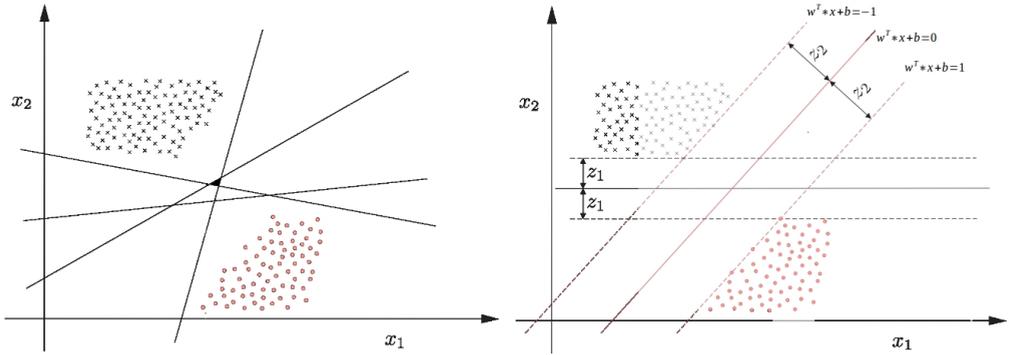
Figure 3.2: Margins in hyperplanes. Image on the left shows the 2 classes, with 2 features x_1 and x_2. Infinite number of lines which can split the population could serve the purpose of separation hyperplane. The image on the right demonstrates the concept of marginal maximization. z_1, z_2 represent the distances between the populations and each decision hyperplane (line). SVM is looking for the hyperplane with maximum z distance. Adapted and adjusted from [87] upon authors approval.

if $w^T x + b > 0$ then $y_i = 1$ and when $w^T x + b < 0$, $y_i = -1$.

The margin to be maximized is the distance between the two lines and is equal to $2z_2$ as shown in figure 3.2. Take any point x_u that lies on the line satisfying the following equation $w^T x + b = -1$. The distance that needs to be maximized is the length of the perpendicular to this line vector starting from x_u and ending on the point x_d that lies on the line described by $w^T x + b = 1$. The point x_d can also be expressed as $x_d = x_u + \lambda w$ (1), where $\lambda ||w||$ is the margin length and $||.||$ denotes the Euclidean distance.

Solving for λ:

$$w^T x_d + b = 1, \text{ given (1) } \Rightarrow$$

$$w^T (x_u + \lambda w) + b = 1 \Rightarrow w^T x_u + \lambda w^T w + b = 1$$

$$\text{but we know that: } w^T x_u + b = -1, \text{ therefore}$$

$$-1 + \lambda w^T w = 1 \Rightarrow \lambda w^T w = 2 \Rightarrow$$

$$\lambda ||w||^2 = 2 \Rightarrow \lambda = \frac{2}{||w||^2}$$

From the last relation and without loss of generality maximizing the distance -$\lambda ||w||$- is equivalent to maximizing $\frac{2}{||w||}$ or $\frac{2}{\sqrt{w^T w}}$ or minimization of $\frac{w^T w}{2}$. Finally, the problem can be expressed as:

$$min_{w,b} \frac{w^T w}{2} \qquad \text{so that} \qquad y_i(w^T x_i + b) \geq 1$$

In practice, several cases involve data that cannot be perfectly separated so the algorithm has to tolerate some misclassified samples. For that a slack variable $\epsilon_i \in \mathbb{R}^p$ for each x_i is used to make the soft-margin extension that has the following form:

$$min_{w,b} \frac{w^T w}{2} + C \sum_{i=1}^{n} \epsilon_i \qquad \text{subject to} \qquad y_i(w^T x_i + b) \geq 1 - \epsilon_i$$

where $\epsilon_i \geq 0, i = 1, ..., n$ and $C > 0$ is the upper bound. The dual form of this is:

$$min_a \frac{1}{2} \sum_{i,j=1}^{n} \alpha_j (y_i y_j x_i \cdot x_j) \alpha_j - \sum_{i,j=1}^{n} \alpha_i \alpha_j \qquad \text{subject to} \qquad \sum_{i=1}^{n} y_i \alpha_i = 0$$

where $0 \leq \alpha_i \leq C$.

The maximization of margins is a constrained optimization problem that can be solved using Lagrange multipliers (for further reading [86–88]). Finally, the decision function of the classifier for a new sample x is:

$$y = sign(w^T x + b) = sign(\sum_{i=1}^{n} y_i a_i (x_i x) + b)$$

It is noteworthy that the maximization of the margin of an SVM is a quadratic function and thus a convex optimization problem, so there is an optimal solution and cannot be "trapped" in local maximum or minimum.

Regularization parameter

A parameter that can be set by the user and influences the final decision boundary is the C parameter, also known as the regularization parameter. It is used to prevent overfitting or underfitting by directly influencing the training loss [89]. The value of the C parameter allows the user to decide how much the algorithm will tolerate misclassified samples during training and thus influence how broad or narrow the margin will be 3.3. Although is not ideal, a common practice in neuroimaging, due to high dimensionality in comparison to the data samples, is to set $C = 1$ without performing further tests [90–92].

Kernels

From the mathematical expression of the decision function of SVM it is clear that the decision of the classifier is defined in terms of inner products $K(x_i, x)$ in input space. Therefore, the problem can be transferred in some expanded feature space by replacing this inner product with another "similarity" function.

$$K(x_i, x) \Rightarrow \Phi(x_i)^T \Phi(x)$$
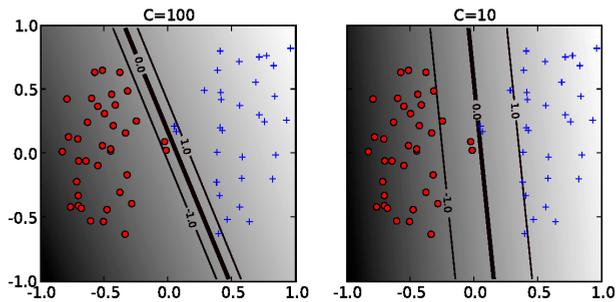
Figure 3.3: The regularization parameter C is used to determine the tolerance level of misclassified samples during the training. Big values of C can lead to an overfitting and therefore to a worse generalization performance on unseen data. On the other hand a very small value of C could underfit the classifier. Adapted from [93]

where K is called a *Kernel* and Φ is called its feature map [94].

Changing the feature space by replacing the similarity function is called the "kernel trick" [95] and allows to separate classes where data are not linearly separable. Popular Kernels are the Radial Basis Function, Polynomial and Sigmoid. By using Kernels [85] the relation between features of samples can be modeled by more complicated, non linear ways. Additionally, with Kernels the classification is performed between paired similarities of the samples, which results in having as many features as samples. When in most cases of functional neuroimaging analysis the datasets are small compared to the number of features, the kernel trick also results to a dimensionality reduction.

The ideal Kernel for each problem is related to the type of the data and the number of features in comparison to the size of the dataset. In practice, a common way to decide which Kernel to use is by testing different Kernels, using cross-validation. However, in neuroimaging the linear *Kernel* is often preferred due to the high dimentionality of data in comparison to the number of samples.

3.2.2 Extremely randomized trees

Another very popular classifier is the Extremely randomized trees (Xtrees) [96]. Xtrees is an extended and developed form of Random Forest Classifier (RFC) and both classifiers are based in the idea of Decision Trees (DT).

Decision Trees

DT are sequential models in the shape of tree diagrams. Such a diagram combines a sequence of different tests in its attribute nodes, to determine a course of action linked to a branch or a sub-tree [97]. In each node, an input value is compared to a threshold value computed from a given training

set. In a classification problem of two classes, one node will split in two and the tree is described as binary tree. Using a binary decision in each node the feature space is partitioned into hyper-rectangles as shown in image 3.4 [87].
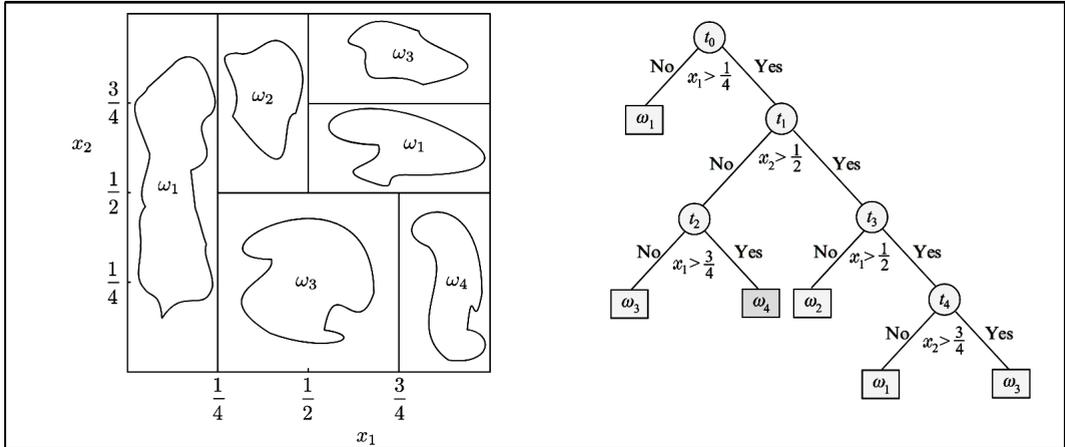
Figure 3.4: Classification Trees are *multistage* systems which split a feature space into hyper-rectangles by performing a comparison to a threshold value, called "splitting criterion", in each node. The left side of the figure demonstrates the partitioned 2-dimensional space for 4 classes, as occurred from the tree shown in the right side. Adapted with author's approval from [87].

Three main issues are raised when designing a DT: What variables are going to be examined in each node, what values will be set as thresholds and how many nodes are needed in each tree. It is best to start answering by defining the objective of each node. For a given dataset X_t that reaches a node t of a binary tree, the splitting criterion aims to split X_t in two sub-sets X_{tL} and X_{tR} (L for left and R for right coming from the direction of the nodes in the visual representation of a node) where the following is true:

$$X_{tL} \cap X_{tR} = \emptyset \qquad \text{and} \qquad X_{tL} \cup X_{tR} = X_t$$

Impurity of nodes

So, the goal in each node is to split the incoming set in the purest possible sub-sets regarding their class-homogeneity. Several methods exist for the quantification of the "impurity" in each split. *Entropy Index*, is one of the most common and is described as:

$$I_{Entr}(t) = -\sum_{i=1}^{n} P(\omega_i|t) \log_2 P(\omega_i|t)$$

where ω_i is class i and $P(\omega_i|t) = \frac{\text{number of samples in } X_t \text{ belonging to } \omega_i}{\text{total number of samples in } X_t}$ known also as frequency of occurrences.

Another popular index of impurity is *Gini Index* expressed as:

$$I_{Gini}(t) = -\sum_{i=1}^{n} P(\omega_i|t)(1 - P(\omega_i|t))$$

Both indexes approach zero if one of the probability values are close to 1 or 0 and they maximize the impurity near 0.5 which is the equiprobable level for the two classes as shown in figure 3.5. Entropy and Gini Indexes produce very similar results [98]. Using the aforementioned indexes, the impurity decrement in a node after splitting the incoming data into the two branches is calculated as:

$$\Delta I(t) = I(t) - \frac{N_{t_L}}{N_t} I(t_L) - \frac{N_{t_R}}{N_t} I(t_R)$$

where N_{t_L} and N_{t_R} are the numbers of samples following the right and the left branch respectively and N_t the total number of samples. $I(t_L)$ and $I(t_R)$ are the impurity values of the nodes created by the two new sets. In the end, the goal is to select the feature with the threshold value that maximizes $\Delta I(t)$. Measuring the impurity provides at the same time an evaluation of the features of the dataset.
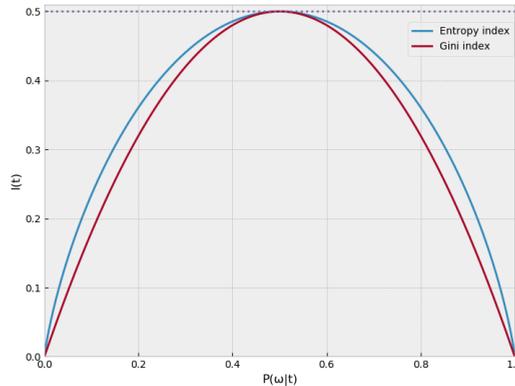
Figure 3.5: Entropy and Gini indexes are used to measure *Impurity* in a node. In x axis is the probability of belonging to one of the classes and in y the corresponding values of Entropy (normalized) and Gini index. When the certainty about belonging to one of the two classes increases then impurity decreases and vise versa.

Stopping and classifying criteria

To stop growing a DT one option is to set a threshold value for $\Delta I(t)$ and decide whether the impurity of the classes does not sufficiently improve further. Another criteria is to set a threshold number in the cardinality of the sets produced in a node. When a node is called *Leaf Node* then there is no other node following. Data that end up at a *Leaf Node* will be assigned with the label of the majority of samples.

Significant advantages of DT which make them attractive to scientists are [99]:

– the interpretability of the results

– the fact that they can handle ordered, categorical or mixed variables/features

– they effectively deal with missing variables

– the existence of an intrinsic feature selection mechanism

– the robustness to outliers

– they do not need any a-priori assumptions to model complex relations between inputs and outputs

Besides the important advantages of DTs, they also suffer from overfitting resulting in some cases to an insufficient generalization on new data.

Ensemble of Decision Trees

A solution to this drawback based on the idea of random subspace learning [100] is to combine many small trees. An ensemble method where many weak classifiers come together to form a strong one. An implementation called RFC [101] is a collection of DTs each one trained in a randomly selected sub-group of the available samples with the method of *Bootstrap Aggregation* or *Bagging* [102]. In order to create more un-correlated predictors and thus reduce bias, a sub-set of the total features is selected with the method of *Bootstrapping* for each tree. In the end, each tree votes equally for one class and the majority decides about the final classification. It is proven that by increasing the number of trees in a RFC is not leading to overfitting [101]. Another ensemble method based on DTs is the Xtrees. Xtrees differs from RFC in the fact that it does not bootstrap the observation in order to built the trees, but rather samples without replacement. Additionally, in Xtrees the nodes are split based on random splits among a random subset of the features selected at every node. Both methods offer several mechanisms for assessing the importance of the features, and thus enhance the interpretability of the model. An advantage of Xtrees over the RFC is that they provide more accurate feature ranking

[99]. RFC often mask the importance of some features, due to the guided structure of the trees. This can lead to a over- or underestimated features. The randomness in splitting of the Xtrees prevent this effect.

Bootstrapping and bagging

Bootstrapping refers to resampling from a given dataset with replacement mostly used for making inferences about a population. When treating sub-samples as surrogate populations and then averaging the measured parameters of those sub-samples, the corresponding real values of the whole population can be approximated [103]. Bootstrapping can also be used for estimating confidence intervals, for testing a hypothesis and for performance estimates. Bagging (Bootstrap Aggregating) is a machine learning technique based on bootstrapping used to improve accuracy and stability of a classification or regression. It is embedded in the algorithm of RFC [102] but its use has been extended further to many techniques and applications. In bagging, given a data set S of size n, i new datasets S_i are being created by sampling from S with replacement and used in different models. *Replacement* permits some samples to exist more than once in an S_i. Each new dataset is used to train a model and in the case of RFC each new dataset trains a DT. Finally, all trained models form an ensemble classifier where a majority vote makes the final decision.

3.3 Estimating Classification performance

3.3.1 Data partitioning

To evaluate the performance of a classifier some predictions need to be performed on an "unseen" part of the data in order to avoid circular processes [104], which result in an overestimate performance. Therefore, as a general rule, the available dataset has to be split in two groups, one for fitting the model and another one for testing the performance. Due to the fact that Neuroimaging datasets are usually small, splitting the dataset is not an option and thus alternative techniques, such as Cross Validation (CV), have to be adopted. In CV the data are split in K sub-sets, where K is an integer. When $K = $ *number-of-samples*, the process is known as Leave-One-Out CV. Each time, one of the K sub-sets is kept for testing and the other $K - 1$ ones are used to fit the classifier. This process is repeated K-times until all parts have been tested. The average of the performances of all repetitions make the overall performance estimation. A particular case is K-fold which occurs when all splits are of equal size. For more accurate estimates, it is recommended to repeat the K-fold process by

shuffling the samples in each fold [105–107]. In case that feature selection process or selection of model parameters of the classification model are involved in the pipeline, then the dataset has to be split in three sub-sets [108] or perform a nested-CV [109].

3.3.2 Confusion Matrix

A decision of a classifier about a tested sample can fall in one of the four following categories:

– *True Positive (TP)*, eg. MCS patients correctly classified.

– *True Negative (TN)*, eg. VS/UWS patients correctly classified.

– *False Positive (FP)*, eg. VS/UWS patients misclassified.

– *False Negative (FN)*, eg. MCS patients misclassified.

A visualization of the counts of these outputs constitute the confusion matrix and is shown in 3.6.

		Real Condition	
		Positive	Negative
Predicted condition	Positive	*True Positives*	*False Positive*
	Negative	*False Negative*	*True Negative*

Figure 3.6: **Confusion Matrix** demonstrates the results of a classification process. It contains all possible outcomes of a classifier. The two classes are described as positive or negative, which is a way to describe the occurrence or not of an effect. If the predicted condition is in accordance with the real one it is described as "True". In a missed classification case there is a "False" instance. From those four numbers many different metrics can be calculated, each one unveiling a different aspect of the classifier.

The aforementioned possible outputs can be combined in several ways and produce different metrics, which estimate the performance of a classifier. Different metrics are useful as in real-life cases a FN can often have more significant impact than a FP and vise versa and that has to be taken under consideration when evaluating the performance.

3.3.3 Metrics

Accuracy is the number of correctly classified samples over the total number of samples and considers misclasified samples from both classes equally. It does not take into account cardinality imbalance of

the classified classes and is calculated as:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}$$

Sensitivity or True Positive Rate or Recall is the probability that a positive sample will be classified correctly and is expressed as:

$$Sensitivity = \frac{TP}{TP + FN}$$

Specificity or True Negative Rate is the probability that a negative sample will be classified correctly and is expressed as:

$$Specificity = \frac{TN}{TN + FP}$$

Balanced Accuracy takes into account class populations. It is calculated as the mean of Sensitivity and Specificity:

$$BAcc = \frac{\frac{TP}{TP+FN} + \frac{TN}{TN+FP}}{2}$$

Precision or Positive Predictive Value measures how accurate the positive predictions are as it takes into account all the samples predicted as positive:

$$Precision = \frac{TP}{TP + FP}$$

F1 score is the *harmonic mean* of Precision and Recall:

$$F1 = \left(\frac{Recall^{-1} + Precision^{-1}}{2} \right)^{-1} = 2 * \frac{Recall * Precision}{Recall + Precision}$$

Receiver Operating Characteristic (ROC) curves are also a very common technique which visualizes the performance of a classifier. It is a plot of the Sensitivities and the Specificities. ROC curves are very informative and besides evaluation of a model, they can be used to decide about parameter values or to compare models [110]. The area under a ROC curve, known as *Area Under the ROC Curve (AUC)*, is often used to evaluate performance. It takes values from 0.5 for chance level performance, to 1 for perfect classification. It expresses the probability that a randomly chosen positive sample will be assigned with a higher probability to be a positive rather than a negative one [111].

3.3.4 Significance of Performance and confidence intervals

In order for the estimated performance of a classifier to be considered as accurate, it has to be tested if it is significantly above chance level. Such evaluation shows whether the measured performance was a matter of chance or not and how confident one can be for the measured performance. When an

independent test set is used, then the Binomial Test is an appropriate method to compare the results of the classifier with the chance level classification [112–114]. When CV schemes are used, tests such as Permutation Test [114] or Paired-comparisons with a Dummy classifier in a bootstrap scheme [115] are more appropriate.

3.4 Features and labels

Features and labels are key terms of ML. Features are the input variables of the classifier. A label indicates to which category belongs a sample in given dataset. Thus, labels are also the output of the decision function.

3.4.1 Features

Features from DOC Neuroimaging data

Though both fMRI and PET consist of voxels, they can produce different kind of features. Regarding FDG-PET in DOC patients, features are extracted from one volume which represents the accumulation of glucose during a specific time-frame, after the injection of the radio-isotope. Therefore, commonly used features extracted from PET are signal at single voxels [116] or averaged over regions of voxels. Accordingly, fMRI is always coming as a sequence of volumes. A very common feature type extracted from fMRI is the strength of functional connectivity, which is actually the degree of temporal correlation between brain regions [43, 117–119].

Feature selection

Not all features are important for differentiating two populations. Some are more important than others and there are features that have no contribution at all. Feature selections methods aim at minimizing the size of the set without decreasing the classification accuracy and preserve the distributions of the classes [120] by removing noisy, irrelevant and redundant features [121]. The benefits of feature selection can be the reduction of computational time and storage demands, improvement of accuracy, facilitation of data comprehension. Feature selection is an open problem and thus there is no optimal technique for selecting the best combination of features in a given classification algorithm. It is noteworthy that sometimes features that seem to be useless by themselves can be of significant importance when they are combined with other features [122]. The only way to make sure of the best subset is to test all possible combinations. Evaluating features by examining the performance of a classifier in all possible combinations is known as *Wrapper* method. Though it is a very accurate way, it cannot be

applied because of the huge amount of feature combinations that occur from high dimensional data. For N features there are $2^N - 1$ combinations and thus it is a NP-hard problem [123, 124]. *Filter* methods perform the feature selection before classification in a test set, using some ranking criteria. Such criteria can be F-score, information gain, mutual information, chi-square. Feature evaluation can also be part of the classification algorithm in the training process, as mentioned earlier in the paragraph of *Decision Trees*.

3.4.2 Labels in the classification of DOC subjects

It is very important in classification to use a clear datasets for fitting the models. Using data with a lot of noise or big heterogeneity under the same label can have a negative impact in the final performance. This problem is a common issue in the field of DOC. The variability of pathological causes which can lead to the same DOC level, different etiologies such as traumatic or non-traumatic, variations in time since the onset result in an heterogeneous population [2]. This heterogeneity complicates the definition of a common diagnostic pattern. Moreover, there is an increased chance of using contaminated labels in the training sample, coming from behaviorally misdiagnosed patients [125]. All the previous reasons result in groups with big variance, which can significantly affect the decision rules of a classifier and set an upper threshold to the performance of the algorithm.

The establishment of a "gold standard" in the characterization of DOC patients is of crucial importance. Possible labels can be the outcome of behavioral assessments or different modalities or the combination of them. Then, performance is measured as the agreement between the classifier and the behavioral assessment [115] or on the diagnosis based on the selected modality.

3.5 Challenges in DOC patients' classification

Besides potential misdiagnosed DOC patients which result to mislabeled subjects in the data fitting a model, there are further parameters that can influence the classification process and jeopardize the model. Available datasets are small in number compared to the size of features, making it one additional parameter that can significantly affect DOC classification. Scans, both PET and fMRI, include tens of thousands of features. This number is substantially higher compared to the dataset size which is so far some tens or some hundreds of subjects. The dataset size often decreases due to subjects that are discarded for reasons like medical treatment, previous neurological disorders and in general not

fulfilling specific inclusion criteria. Additionally, subjects are often rejected due to big deformations in the brain that cannot always be handled properly by the preprocessing algorithms. Moreover, some brain mechanisms, such as plasticity have not been modelled and as such cannot be imported or be taken under consideration by the known classification algorithms. Spasticity and lack of communication with patients also increases scanning noise and affect classification.

3.6 Summary

ML consist of complex algorithms, which combine many scientific fields to provide predictions in new data. In neurosciences they are used to form diagnostic or prognostic tools as well as to uncover structural or functional patterns of the brain. A variety of metrics and of data splitting methods are used to express different purposes and minimize pitfalls' risk, which can be vital for patients and of significant ethical importance. In the field of DOC not much research has been undertaken so far in the analysis of FN using ML. The latter is a result of data and computational limitations which, with the time passing, are being solved. The next chapters demonstrate applications of ML in DOC.

Chapter 4

Differentiating low levels of consciousness with intrinsic functional connectivity

Demertzi A*, **Antonopoulos G***, Heine L, Voss H, Crone J, de Los Angeles C, Bahri M, Di Perri C, Vanhaudenhuyse A, Charland-Verville V, Kronbichler M, Trinka E, Phillips C, Gomez F, Tshibanda L, Soddu A, Schiff N, Whitfield-Gabrieli S, Laureys S.

[*] Contributed equally

Summary

Statistical analysis findings in resting state connectivity remain challenging for clinicians to use at the single-subject level. In this study a classifier separating minimally conscious from unresponsive patients, it is presented. Functional connectivity networks from fMRI scans acquired in three different centers were extracted, to evaluate their discriminatory power in differentiating MCS from VS/UWS patients. Each network was tested for its contribution to consciousness, using feature evaluation methods. Additionally, the estimated classification model was tested, using five different datasets to search for its generalization capacity and the border line conditions. Results yield very good classification performance between MCS and VS/UWS patients without sedation. Generalization with other datasets shown that the classifier is capable to differentiate between impaired consciousness and pathological lack of consciousness.

Keywords*: Functional Connectivity, DMN, SVM, Feature Selection, Classification, MRI- fMRI, DOC*

4.1   Introduction

Despite the significant progress in functional connectivity studies in DOC, it is still challenging for clinicians to apply those findings and assist diagnosis of the level of consciousness at the single-subject level. Up to now, an accurate diagnosis of patients has been achieved with the use of Trancranial Magnetic Stimulation (TMS) in combination with EEG [126, 127], later by using only EEG features [128, 129] and in terms of glucose consumption and PET by using the brain hemisphere with the highest glucose consumption [50]. Regarding fMRI studies, active paradigms [52, 130] and passive paradigms in terms of verbal and tactile stimulation studies [131, 132] have been performed aiming at identifying alterations in the activity of each functional network. In rsfMRI there is no research so far that could provide an assisting diagnostic tool to clinicians. The rsfMRI provides temporal and better spatial resolution compared to PET. In addition, it does not have negative harmful radiation and does not require for a tracer to be injected into the subjects. MRI scanning is a detrimental examination for brain lesioned patients and it only takes an additional 10 minutes of scanning to acquire the functional scans. In total, MRI demonstrates significant advantages over other techniques which make important merits of rsfMRI. The challenge, though, remains on how to make sense of the acquired signal and how to use it efficiently.

Objectives

In this study we investigated whether a system-level brain organization can provide insights in order to diagnose DOC patients individually. More precisely, the performed research aimed at evaluating the contribution of functional networks to the level of consciousness and investigating if there can be a prediction model that can translate the network contributions to a diagnostic tool.

4.2   Subjects

Three datasets were used to train, validate and test the model. 51 patients were scanned in Liege without sedation, from which 26 were in MCS, 19 in VS/UWS and 6 in coma. Two datasets of non-sedated DOC patients from collaborators in Salzburg (Department of Psychology and Centre for Neurocognitive Research) and in New York (Department of Radiology and Citigroup Biomedical Imaging Centre of Weill Cornell Medical College), consisting of 10 and 5 MCS and 5 and 1 VS/UWS respectively. One EMCS was also included in the dataset from New York. Additionally, 21 healthy subjects from the Centre Hospitalier Universitaire (CHU) de Liege and 18 from Salzburg, were used. A group of 53 patients (33 MCS

and 20 VS/UWS) under sedation were scanned in Liege. Additionally, we tested 9 healthy controls that were scanned at the Cyclotron Research Centre in Liege, during sedation process with propofol. rsfMRI was acquired during the whole process and thus connectivity values of both awake and deep sedation were available. All sedated subjects were evaluated using the Ramsay Sedation Scale [133]. Finally, 8 congenitally deaf and 9 congenitally blind subjects, without disorders of consciousness from Panum Institute, University of Copenhagen.

For the evaluation of patients' level of consciousness, repeated clinical examinations with the CRS-R (average number of assessments n= 6 per patient) were performed. The clinical diagnosis was further confirmed with FDG-PET imaging, which has been shown to have high sensitivity in identifying patients in MCS [51]. Therefore, patients with an ambiguous profile on clinical assessment and neuroimaging data or those that diagnosis changed within seven days from scanning and subjects with technical issues, noisy images, were excluded from the study.

4.3 Connectivity analysis

Preprocessing

Preprocessing was performed using SPM8 and is demonstrated in steps 1 to 3 on the left column of figure 4.1. Each structural T1 image was used as reference image on which fMRI sequential volumes were co-registered. At first, all T1 scans underwent manual re-orientation to match a template, they were normalized into MNI stereotactic space and then segmented using the "Unified segmentation" of SPM. Sedated patients were spatialy normalized using diffeomorphic anatomical registration through exponentiated Lie algebra (DARTEL) [134], using an average template of patients and healthy controls. Finally, spatial Gaussian smoothing with a 6mm FWHM kernel was applied to all produced images.

Head motion can decrease or increase the measured signal and make it, wrongly, be considered as neuronal activity [135], hence, we performed artifact detection and removal. Volumes were compared to previous ones and if they had a displacement bigger than 2mm in any of the 3 axes, or a rotation bigger than 0.02 rads, or global intensity higher than 3 standard deviations of the global mean intensity of the entire sequence, they were used as nuisance regressors. De-noising was performed using the signal component-based Noise Correction Method (CompCor) [136]. White matter and Cerebrospinal Fluid (CSF) masks were eroded by one voxel to account for partial voluming with the gray matter [137]. To restrict the signal to low frequencies, which are in the range of BOLD [33, 138], band-pass filter
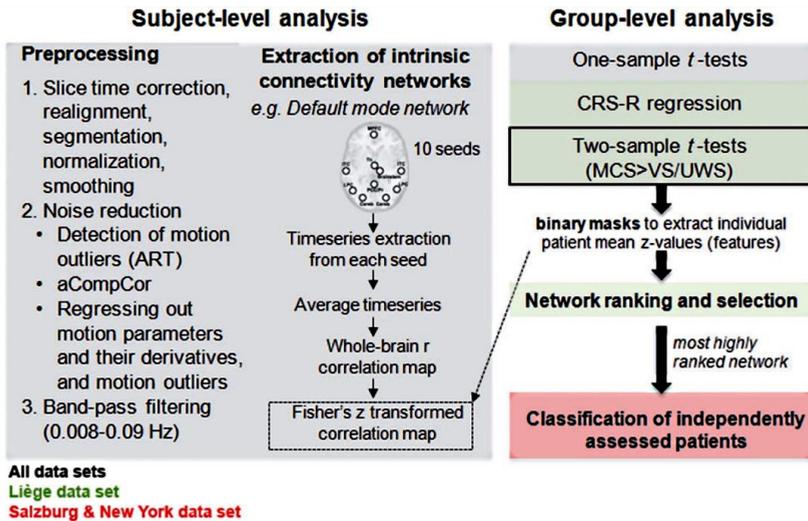
Figure 4.1: Analysis pipeline. The left frame illustrates the single subject level analysis including: image spatial preparation, corrections for movements, de-noising and connectivity networks extraction. The right frame describes the data analysis part of this study. First, statistical contrast between MCS and VS/UWS was calculated for each network from which the binary masks were created. The masks were used to extract the regional connectivity values that were ranked in the next step to estimate the discriminatory capacity of each network. Model evaluation was performed in the last step using "unseen" data. Adapted from the published version.

of 0.008–0.09 Hz was applied on the time series. Finally, we regressed out the residual head motion parameters.

Image preprocessing was performed in SPM8. For Artifact detection and Removal we used Artifact Detection Tools (ART) `www.nitrc.org/projects/artifact_detect`. For denoising and connectivity analysis we occupied methods as incorporated in the CONN toolbox [139].

Connectivity values and networks

To estimate functional connectivity between brain regions, we elaborated a seed-voxel approach. The seed-voxel approach was preferred over Independent Component Analysis (ICA) as the latter -applied to patients which have big lesions- can affect the identified networks. ICA, being data driven, can increase potential need for manual correction or rejection of subjects [117]. In the seed-based approach neuro-pathological network disruption can be dealt using enlarged spherical seeds. We used spheres of 10mm radius for cortical and 4mm for subcortical regions, which were combined accordingly to form the desired networks. Default mode, frontoparietal, salience, auditory, sensorimotor and visual networks were examined in this study. The coordinates of the seeds corresponding to each network

were selected from the literature [35, 140–145].

Time series were extracted from all voxels of each network/Region Of Interest (ROI) and then were averaged. From the averaged time series we estimated whole-brain correlation r-maps which were then converted to normally distributed Fisher's z transformed correlation maps to permit for group-level comparisons.

Networks' Functional connectivity values for MCS, VS/UWS and coma patients were estimated using one-sample t-tests in the dataset of non-sedated patients scanned in Liege. Healthy controls were used to cross-check for the validity of networks' characterization. Aetiology (traumatic/non traumatic) and chronicity (acute/chronic) were included in a 2x2 factorial design to test interaction effect with clinical evaluation of patients (MCS and VS/UWS) and variables were included as regressors in the general linear model, if effect was present.

Linear regression was applied to CRS-R scores and the connectivity values of each network, to examine the contribution of each network in the level of consciousness. The same process was applied for the cerebellar network. As it is known for not being associated to consciousness related processes [146, 147], it was utilized to control the contribution to the level of consciousness. The cerebellar network was extracted using seed regions on the posterior lobe bilaterally [144] and on the inferior semi-lunar lobule [140].

4.3.1 Features' extraction and evaluation

A t-test between MCS and VS/UWS, for each network, provided the contrast maps which then were thresholded to be utilized as masks. The masks were used to extract the mean connectivity of each network (average z-values across the whole mask) from each subject, resulting in 6 values per subject. The extracted values served the purpose of features and were tested for their discriminatory power with two feature-evaluation methods [148]. At first, feature ranking using a t-test implementation (`http://www.mathworks.nl/help/bioinfo/ref/rankfeatures.html`)[149] and then, a single-feature classification with SVM [123] were performed.

4.3.2 Classification of DOC patients

In order to automatically classify patients, we focused on the network which was ranked higher in the evaluation process. We used a linear SVM with default regularization parameter C=1 [92, 150]. The training of the classifier was performed using the Liege dataset and for testing we used the datasets

from New York and Salzburg. In all cases we used SVM, we used the libsvm [151] as implemented in matlab.

4.4   Results

4.4.1   Networks involvement in consciousness

Controls' functional connectivity for all networks was identified in all regions as it is classically reported in the literature. Connectivity values appeared to be smaller in MCS patients compared to controls. Connectivity in VS/UWS was weak and in some cases hardly identifiable. The correlation analysis that was performed between network connectivity and CRS-R scores, demonstrated good correlation between them, as illustrated in figure 4.2. The cerebelar network, as expected, did not correlate to the behavioural evaluation.
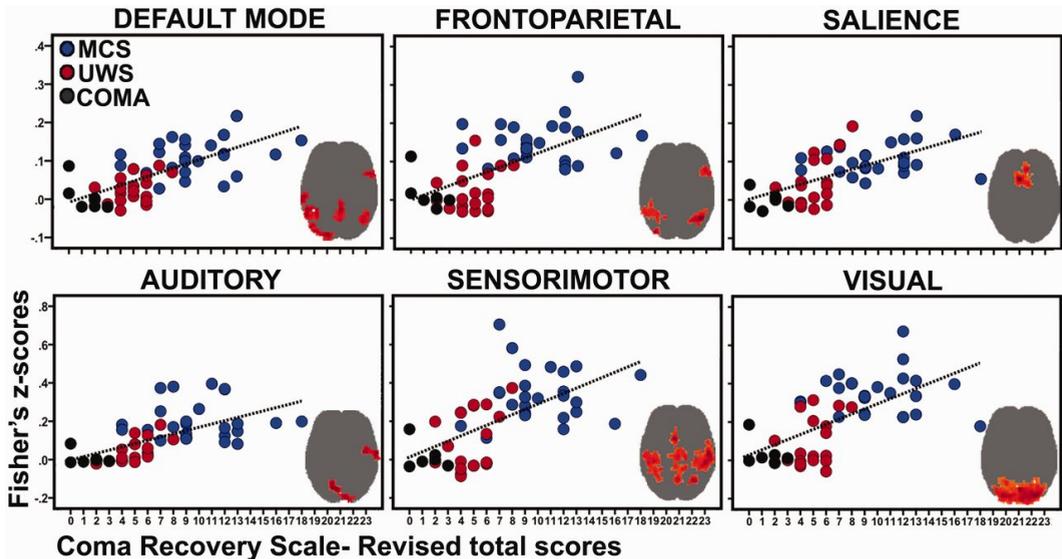
Figure 4.2: Correlation between CRS-R scores and functional connectivity values. Regions in red represent the areas from which connectivity values were extracted. Functional connectivity correlates with behavioural scale evaluation. Statistical maps are thresholded at FWE P < 0.05 (cluster-level)

Contrast maps, from which network masks were extracted, are illustrated in figure 4.3. To minimize the possibility that differences in functional connectivity reflected differences in brain anatomy, we performed a two-sample t-test voxel- based morphometry on the normalized gray matter and white matter segmented masks (smoothed at 6mm full- width at half-maximum). No differences in gray matter volume between patients in MCS and VS/UWS were identified at FWE $P < 0.05$ either at the whole-brain

or at the cluster-level. Similarly, the analysis of white matter volumes showed no differences between the two groups, even at a liberal threshold $P < 0.001$ (whole brain level) uncorrected for multiple comparisons.
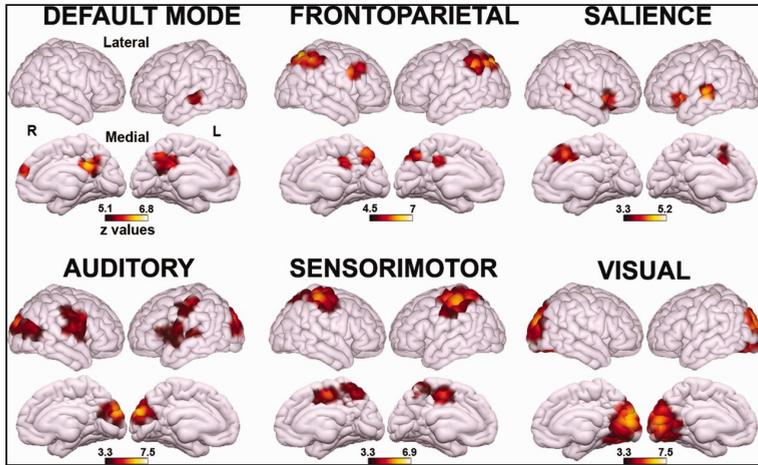
Figure 4.3: Contrast maps between MCS and VS/UWS patients. Regions showing higher functional connectivity in MCS patients compared to patients in VS/UWS for each network. Statistical maps are thresholded at FWE P < 0.05 (cluster-level) and are rendered on 3D surface plot template (top = lateral view; bottom = medial view).

4.4.2 Feature ranking results

In both methods for feature evaluation, the Auditory Network was found to have the highest discriminatory power. Visual, DMN, Frontoparietal, Salience and Sensorimotor were following in this order when evaluated with the t-test method. Salience was second in the single value classification only by one TP better than DMN, Frontoparietal and Salience. Detailed results can be found in table 4.1.

Network	Feature selection criterion (t-test)			Single-feature classification		
	t value	Rank	p value	True positives (MCS)	True negatives (VS/UWS)	Performance accuracy
Auditory	8.32	1	<.001	25	18	43/45
Visual	7.79	2	<.001	23	15	38/45
Default mode	6.95	3	<.001	23	15	38/45
Frontoparietal	6.82	4	<.001	23	15	38/45
Salience	6.21	5	<.001	24	15	39/45
Sensorimotor	5.87	6	<.001	24	13	37/45

Table 4.1: Evaluation of networks using feature evaluation methods. Results of feature ranking based on t-test on the left of the table and Single-Feature SVM classification results on the right side of the table. For both techniques, the Auditory network provided better discriminatory power between MCS and VS/UWS

4.4.3 DOC patients classification

To avoid classifying with only one feature and possibly overfit the model, we extracted values from bilateral auditory and visual cortices, that are the clusters that the Auditory Network consists of. From the 22 tested subjects, two were not classified correctly, one from each group. We included the EMCS patient with the MCS to test how would it be treated by the classifier, and it was classified correctly. The confusion matrix with the classification results is shown in the following table 4.2.

		Label	
		MCS	UWS
Classifier	MCS	15	1
	UWS	1	5

Sens = 15/16   Spec = 5/6
Acc = 20/22

Table 4.2: Confusion matrix of the independent dataset. One false negative and one false positive yielding to a Specificity of 83 % and a Sensitivity of 94 %. The total Accuracy is 91%

An illustration of the separation hyperplane and the data can be found in (figure 4.4).
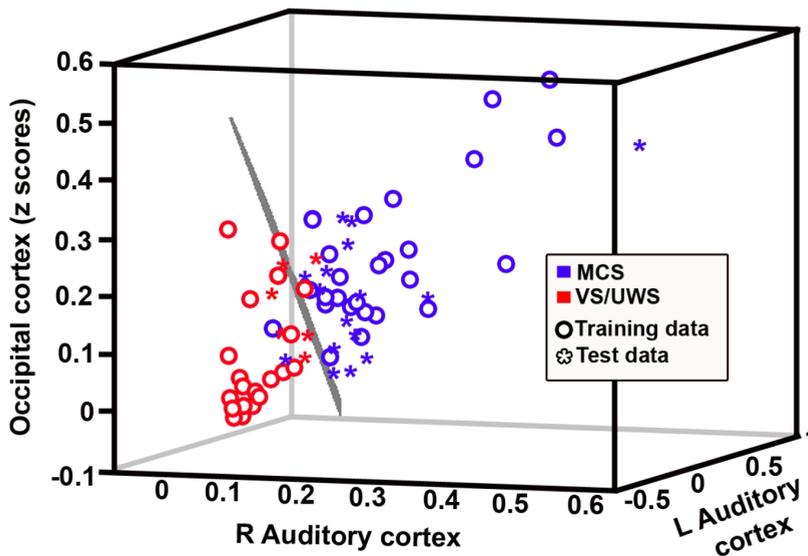
Figure 4.4: Classification Results. Three dimensional representation of the classification. Axes represent the values of the three main clusters of the Auditory network: Bilateral auditory cortices and occipital cortex. The hyperplane is calculated by training with the Liege dataset set. One patient from each group was misclassified resulting to an accuracy of 91% (20/22).

4.5 Generalization of the classifier to a wider population

The classification between the two DOC subgroups, MCS and VS/UWS, is an equivalent of finding a boundary between subjects with residuals of consciousness and without consciousness. In other words, the classifier decides between the appearance or the absence of consciousness. To get an idea of what the classifier is measuring and what it is not, generalization trials had to be performed to investigate boundary and precarious conditions [108].

4.5.1 Tracking consciousness

A model able to detect signs of consciousness has to outperform in controls. The two healthy groups mentioned earlier, from Liege and Salzburg, were used to confirm that the classifier is indeed capable to capture consciousness. We tried the classifier on these groups and 2 out of 39 subjects were identified as non-conscious. Both misclassified subjects were relatively close to the decision boundary, as illustrated in figure 4.5.
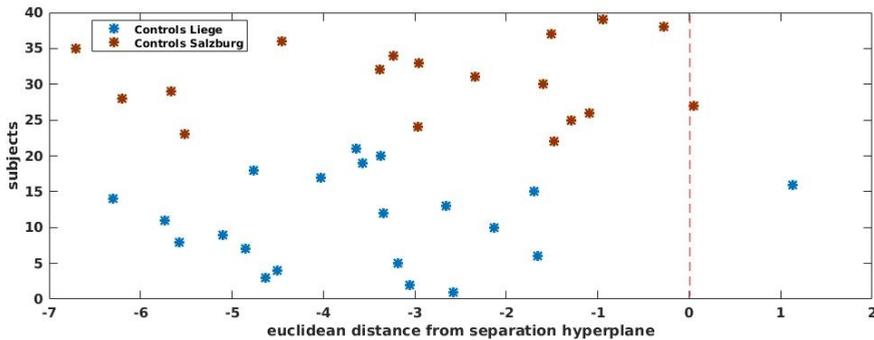
Figure 4.5: Classification of controls in a 2D representation using the Euclidean distance from the separation hyperplane. 2 out of 39 healthy subjects (95%) are placed in the consciousness side showing a high True positive rate.

4.5.2 Classifier's evaluation on healthy, sedated subjects

To take generalization of the classifier a step further we tested the classifier in healthy sedated population. Connectivity values during wakefulness and during deep sedation were extracted from each subject. Though all subjects before sedation were classified as "conscious", during the deep sedation classifier detected 5 of them as conscious and the rest 4 as non conscious. The two dimensional illustration of the results are in figure 4.6.
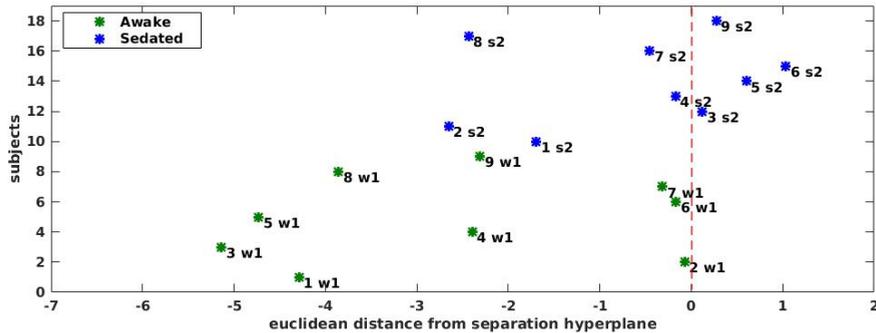
Figure 4.6: Classification results of healthy subjects under wakefulness stage 1 (w1) and deep propofol anesthesia stage 2 (s2) sedation. In a 2D representation using the Euclidean distance from the separation hyperplane, five deeply sedated subjects are placed in the classifier in the side of "consciousness".

4.5.3   Testing the classifier in congenitally deaf and blind subjects

The identification of Auditory network, that encompassed temporal and occipital regions, as the one with the best discriminatory capacity is raising the question whether the classifier captures a higher-order cortical organization pertaining to conscious conditions or merely a sensory function (i.e. auditory, visual). To clarify that, we tested the classifier with one group of conscious but congenitally deaf subjects and with one of conscious but congenitally blind subjects.

Out of the nine congenitally blind patients one has been detected as not conscious from the classifier leading to an accuracy of 88.9% as shown in figure 4.7a. From the 8 congenitally deaf subjects one was classified as non conscious. A 2-dimensional representation of the classification results can be found in figure 4.7b.

4.5.4   Classification of sedated patients

The last dataset we used to test the classifier was part of the sedated patients. Specifically, 53 sedated patients were tested, from which 33 were in MCS and 20 in VS/UWS. Results were similar to those of sedated healthy subjects. Near-chance level classification was achieved, with 35 out 53 patients being classified correctly (66%) 4.8. An overview of the results can be found in table 4.3

4.6   Discussion

In this research we investigated how functional connectivity networks can be used to differentiate patients with DOC. At first we evaluated how much the connectivity strength of networks is in accordance
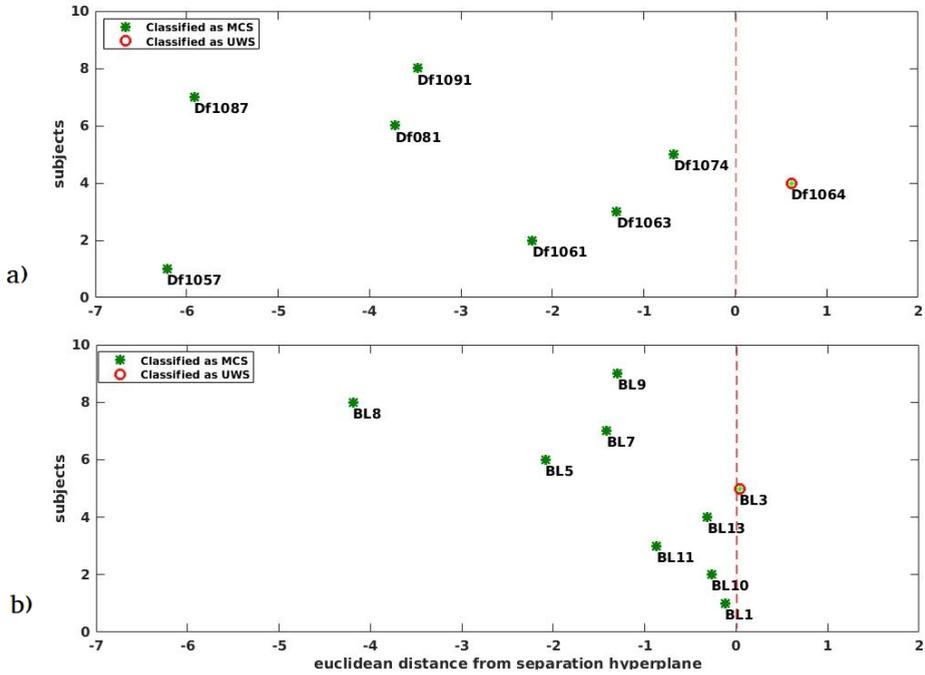
Figure 4.7: Classification results of congenitally deaf and blind subjects in a 2D representation using the Euclidean distance from the separation hyperplane. a) Seven subjects out of eight were classified as having signs of consciousness. b) One congenitally blind subject out of nine was detected as unconscious.

		Label	
		MCS(33)	UWS(20)
Classifier	MCS	24	9
	UWS	9	11

Sens = 24/33   Spec =9/20

Acc = 35/53

Table 4.3: Confusion matrix of the sedated patients' dataset. Results give Sensitivity of 73%, Specificity 45% and AUC of 66%.
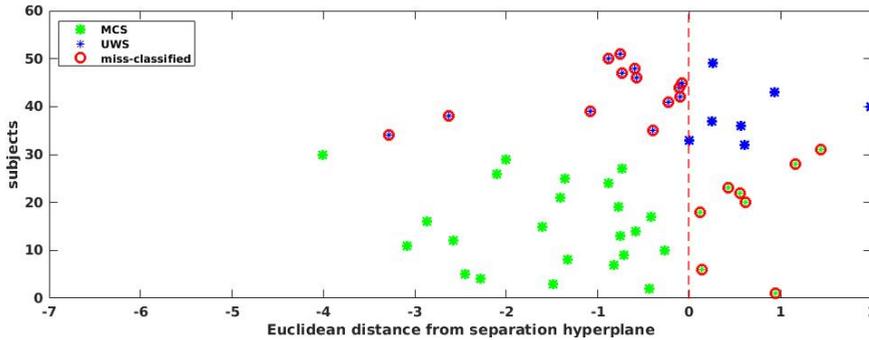
Figure 4.8: Classification results of sedated patients. All misclassified subjects are annotated with the red "o". There is a big overlap between the two populations which explains the chance level classification.

with the CRS-R scale and we found a strong correlation between them. As feature selection methods do not have a theoretical frame for optimal methods, we used two techniques from two different categories, one from *filter* methods and one from *wrapper*. Both methods highlighted the Auditory network as the best one for classifying patients. To restrict the phenomenon of contaminated labels, which can significantly affect the decision boundary and thus the classification results, we only included subjects for which the clinical evaluation was in accordance with the PET evaluation.

A critical point of this work is the usage of the same dataset for the identification of the differences between the two populations, and also for the evaluation of the features. This yields a circularity in the processes, which can result to over-estimation of the performance. We overcame this issue by validating the classifier with datasets that were not involved in previous steps. The use of these two datasets helps to overcome the issue of double dipping [104, 152]. The fact that we used the same data for creating statistical maps as well as for feature evaluation and selection could lead to an overestimation of the performance, if there were not the two unseen datasets for the final performance estimation of the model.

The unseen datasets demonstrated a very good accuracy. Interestingly, in the misclassified subjects there was one VS/UWS which emerged to higher state a few weeks later and an MCS.

Furthermore, we tested the model using un-sedated and sedated healthy subjects. That way, we wanted to examine whether or not the classifier is indeed able to detect consciousness. All un-sedated controls were classified as conscious, but deeply sedated healthy subjects demonstrated chance level results. It is noteworthy that for two of the subjects the connectivity values of deep sedation appeared increased compared to unsedated-state values.

The classification of healthy controls during sedation aimed at testing the generalization of the classifier in different un-consciousness etiologies. Results confirmed that the classifier is performing well in detecting consciousness as it assigned again all un-sedated subjects as conscious. Nevertheless, the number of misclassified subjects in deep sedation raises questions. Does lack of consciousness have a unique pattern or there might be more ways to be unconscious? The random level classification results of sedated patients point to pharmacological sedation having a different brain connectivity pattern to the one that the classifier is trained to disentangle. Meaning, that the classifier is incapable of assigning non pathological unconsciousness. To address this issue further studies should be performed, in which included sedated patients and controls will be sedated with the *same* drugs, as the sedated patients of our study were under *different* anesthetic drugs.

The accurate classification of the congenitally blind and deaf groups shows that the "Auditory" network is not highlighted due to some sound stimulation coming from the scanner or some visual stimulation. A reason could be the cross-modal interaction of visual and auditory cortex that is considered relevant for multisensory integration [153] and has been suggested as a facilitator for top-down influences of higher-order regions to create predictions of forthcoming sensory events [154]. Additionally, preserved functional MRI activity in temporal and occipital areas has been shown for healthy subjects, who were attentive and aware of the auditory violations, during mental counting of auditory temporal irregularities [155]

In conclusion, we have shown that the classification of patients using functional connectivity provides an important assisting tool for the evaluation of consciousness. Multisensory integration supported by crossmodal connectivity have been highlighted by machine learning methods, as of great importance for consciousness.

The validity and significance of our findings could be further improved when including more data. More data coming from different unconscious populations, such as pharmacological or physiological unconsciousness, could help build a more robust model with a higher generalization capacity.

Classification of DOC patients using standardized glucose metabolism

Summary

In this study we used machine learning for differentiating between unresponsive and minimally conscious patients, using PET scans. We aimed at creating a tool to assist clinicians with the diagnosis of DOC patients. We used SUV of FDG-PET and estimated parameters of a model with three classifiers, trying to minimize misclassification of MCS patients. We used 158 patients in a cross-validation scheme to build the model and 53 subjects to validate it. We achieved a Recall 89%, a Precision of 89% and a general accuracy of 85%.
Keywords*: DOC, FDG, PET, Feature Selection, consciousness classification, SVM, Random Forests*

5.1   Aim of the project

PET has shown to be very robust for diagnosing the level of consciousness in patients suffering from brain lesions, by visual examination of the SPM analysis of hypometabolic and preserved regions [51]. In DOC patients, decreased local and global metabolic activity was reported compared to healthy subjects, varying between 40% and 50% in the acute phase and 30% to 40% in the subacute and chronic phase [46, 47, 156]. Differences in the metabolic activity of the intrinsic midline network and the extrinsic cerebellar network were found between the different levels of consciousness and controls [157]. No difference was reported between the DOC groups.

Although the resolution of PET is not the best among FN techniques, it has the advantages having no noise during scanning and thus it does not distract the subject, and being unaffected by magnetic properties of the tissues. Additionally, the imaged activity is that occuring an earlier time window to the scanning time. This property allows subjects to be sedated during the acquisition while the captured metabolic activity represents a time frame prior to the sedation. This makes PET less prone to motion artifacts, which is important in DOC as patients often suffer from spasticity or perform random movements during scanning. Motion during scanning can result in very noisy and practically useless scans, therefore sedation is often required.

Discriminating MCS patients from VS/UWS ones, using PET, has been performed in previous studies [50, 158]. The first study suggested a correlation of the level of consciousness and the overall cortical energy turnover. Additionally, it defined the threshold of the shift between unresponsive and minimal consciousness to be 50% of healthy controls' metabolic activity. A voxel based analysis highlighted primary sensorimotor areas, adjacent frontoparietal regions and precuneus, being in line with previous studies [156, 159, 160], which had reported those regions as important for consciousness. No validation with unseen data was performed in the study. The second study [50] examined the hypothesis of having a minimal glucose metabolic requirement for a patient to be conscious. This minimum requirement was calculated in the hemisphere of patients that had the highest mean consumption, and was found to be at 42% of healthy subjects' cortical activity. No further analysis was performed to highlight more precisely the most significant regions within the selected hemisphere.

This project aims at differentiating MCS patients from those being in VS/UWS using brain regions defined by anatomy. Additionally, SUV values of PET were preferred in order to permit for inter-scanner generalization. For the classification process, emphasis is given on detecting the presence of responsiveness with high accuracy (here detecting MCS patients). This objective originates from clinical and

ethical requirements which dictate the medical management of patients, balancing the autonomy and well being. Therefore, we aimed at minimizing the amount of MCS patients detected as VS/UWS, hence we evaluated our results in terms of *recall* and *precision*. Additionally, we wanted to investigate the spatial patterns of consciousness using a data driven approach. We assume that some brain regions will be of greater importance than others and moreover, regions from both hemispheres will play a role in the classification of the two groups.

5.2 Datasets

Our dataset consists of PET scans of 211 patients (159 MCS and 70 VS/UWS) and of 20 healthy subjects, free of psychiatric or neurological history. The scan was performed 52±13 minutes after intravenous injection of 150 or 300 MBq of FDG using a Gemini TF PET-CT scanner (Philips Medical Systems). A low-dose CT was acquired for attenuation correction, followed by a 12-minute emission scan. The studies were reconstructed using a LOR-OSEM algorithm and reconstructed images had 2^3 mm isotropic voxels in a 256x256x89 voxel matrix. An examiner was present during the whole acquisition to ensure that the patient remained awake and eyes open in a silent and dark room (tactile or auditory stimuli were administered when patients were closing their eyes). Patients were assessed at variable times after the brain injury (acute<3 months since onset-75 patients-, or chronic stage>3 months since onset -136 patients-), in order to clarify the actual state of consciousness. All patients were repeatedly evaluated with the CRS-R [16]. At least 5 CRS-R assessments were performed by accredited trained specialists. In all cases, assessments were performed on different days and clinical labels were assigned according to the highest clinical diagnosis across all evaluations.

5.3 Methods

The dataset was randomly split in two subsets, where each had the same proportion of MCS and VS/UWS as in the full dataset. More precisely, one subset included 158 patients (106 MCS, 52 VS/UWS) and the other one included 53 patients (35 MCS, 18 VS/UWS). The first and largest subset was used to train and estimate the performance of the classification models (test-set) and the other one to validate the selected models (validation-set).

5.3.1 Standardized Uptake Values and quality check

We created the SUV-PET scans from the DICOM format data with our custom code[1], following the steps described in the guidelines/pseudocode of the Quantitative Imaging Biomarkers Alliance [27].

5.3.2 Preparation of the scans

To ensure that homologous brain regions are compared, all PET images were spatially normalized into MNI stereotactic space. Using the standard PET template of SPM can be problematic for our dataset, because the PET template of SPM is created from ^{15}O-H_2O cerebral blood flow PET scans and only from healthy controls. In our case, the injected tracer is ^{18}F-FDG and also our subjects have very often brains with big deformations. To address those issues, we created a study specific template using only the scans of the test-set. We performed the following steps as described in [116] to create a template targeted to our subjects: all images of the test set were at first normalized with the SPM standard PET template, along with the a-priori white matter and gray matter Tissue Probability Maps (TPM) as implemented in SPM12 [161]. Then, all normalized images were averaged and smoothed with an 8mm FWHM kernel to create the study template. We normalized all original SUV-PET images using only the study template. In the latter normalization procedure the regularization imposed on the nonlinear warping was increased by one order compared to the standard setting to prevent unrealistic warping [162]. Finally, the normalized scans were smoothed with a FWHM = 8mm Gaussian filter. All normalization processes were performed using the *Old Normalization* option of the SPM12 package. The specific normalization algorithm applies an affine registration, followed by estimating nonlinear deformations. The deformations are defined by a linear combination of three dimensional discrete cosine transform basis function.

Parcelation and extraction of features

The latest version of Automated Anatomical Labeling atlas [163, 164], which indicates macroscopic brain structures, was selected for our analysis. The atlas consists of 120 regions: 47 volumes of interest in each hemisphere of the brain, and 26 areas of the cerebellum, which we did not include in our analysis. For each PET scan, we extracted the mean metabolic activity of each region by averaging the included voxels. Each regional mean formed one feature.

---

[1]`https://github.com/antogeo/PET_classification`

All scans were controlled for their quality using z-score with the threshold at 3. Z-scores were applied in the mean values of intracranial activity and in the region of each subject with the highest mean intensity.

5.3.3 Building the classification model

We used scikit-learn [165] for all the steps in testing and validating the models.

Classifiers

We selected two classifiers from different families for our analysis. We chose a linear SVM [86] and a non parametric Xtrees [102], which incorporates a feature evaluation. We wanted to estimate two SVM classifiers with different sensitivities in identifying the two populations. As we aimed at minimizing MCS misclassification, we wanted a classifier to emphasize on high *recall* and the other one in high *precision*. We consider that the hyperplanes of those two classifiers will be parallel or almost parallel to each other in the space of our interest [166]. In the area between them there will be subjects that are classified as MCS by the one classifier and as VS/UWS by the other one. We wanted to test the performance of a Xtrees for this area, as it is partitioning the space in a very different way and thus decides with different criteria. Also, by using the whole feature set in Xtrees we can have a more straight forward interpretetion of the regional contribution in the classification. Moreover, the redundancy inherent to the joint implementation of different classifiers can be a source of final decision improvement [167]. In such a way when the two classifiers make the same prediction for a sample, there is high confidence about the result and when not then the decision will be driven by the Xtrees. Figure 5.1 illustrates the aforementioned concept. Such structure is described as tree-like classifier structure (classification tree) and is a serial strategy to combine decisions with dynamic classifier selection [167]. During the model selection we evaluated the performance of the tested models in terms of AUC, *recall* and *precision*. AUC was calculated in order to provide us with an estimate of the separability of the two classes.

Estimation of the parameters of the models

The selection process of the parameters was performed in the test set (158 subjects). Features were normalized by removing the median from each feature and then scaled with the interquartile range, which is the range between the 1st quartile (25th quantile) and the 3rd quartile (75th quantile). For the linear SVM models, we performed a feature evaluation by ranking the features using ANOVA F-values and we tested for the optimal number of best features in terms AUC and F1-score. The process was
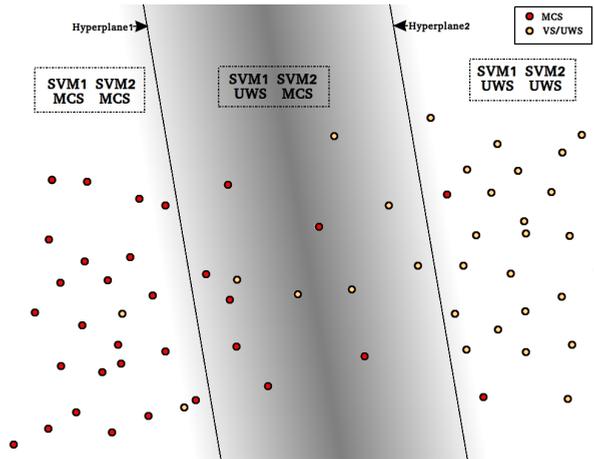
Figure 5.1: The idea of the Classification model in 2D. We assume that the two SVM classifiers will form two parallel decision lines (Hyperplane 1 & 2). The subjects that are on the left side of hyperplane 1 are assigned as MCS and those on the rights side as VS/UWS. Respectively, for hyperplane 2 the subjects that lie on the right side will be classified as VS/UWS and on the left side as MCS. Samples that lie in the middle area are assigned differently from the two SVM classifiers and the Xtrees will make the final decision.

performed by randomly splitting the test set further, using 70% of the data to train and 30% to test. For each number of features we performed 100 differet splits and always kept the same proportion of the two groups as the initial dataset. The Xtrees has an internal feature selection process, thus we did not apply any feature selection prior to classifications.

We wanted to have two SVM classifiers, one with *recall* of 95% and one with *precision* of 90%. Therefore, we tested penalizing with higher values the misclassified samples of one class, over the other class, during the training phase. A range of weights starting from penalizing VS/UWS 10 times less than MCS (1/10) and up to the reverse ratio (10/1) were tested. During the model evaluation we performed 5-fold cross validation, for each value, iterated 50 times, to ensure more accurate estimates of the performance [168]. In all folds of all iterations, the proportion of the two classes was similar to the one of the full dataset. For the Xtrees, we used the same parameters as in [129] which facilitate the feature evaluation, performance and computational complexity.

Statistical analysis

To validate the estimated models and the confidence intervals of our estimations we used the unseen data of the validation set. To compare our results with the empirical chance level we used a *dummy classifier*, which is a model selecting the "most frequent" class for all subjects [165]. We performed

bootstrapping [103] with 1000 repetitions in the validation dataset and classified them using the selected models trained in test set. For each iteration we compared our models to the *dummy classifier* [115]. In other words, we created a distribution from the paired differences between our models and the empirical chance level as expressed by the *dummy classifier*.

5.3.4   Classification of healthy subject

Finally, we classified the group of controls with the three classifiers expecting that they will be treated as MCS subjects.

5.4   Results

5.4.1   Image preparation

All SUV scans of the subjects, produced by our custom code, had z-scores smaller that 3. One control had z-score above than 3 and was excluded from the classification.

The study specific template had bigger ventricles as compared to the template of SPM. The two templates are shown in figure 5.2a). Figure 5.2b) shows the fit of the parcellation atlas on the custom template and two patients from the dataset we used to make the model.
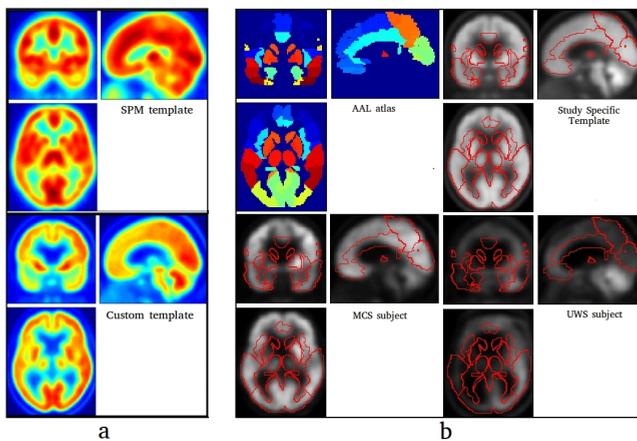
Figure 5.2: a) Study specific template (on the bottom) was preferred over the standard PET template (on the top). The standard template existing in SPM is made out of healthy controls and a different tracer (H_2 ^{15}O). Our custom template has bigger ventricles. **b) Fitting AAL atlas.** Upper left image shows the AAL atlas in "jet" color-map. Upper right shows the study specific template and bottom row shows a MCS subject on the left and an UWS subject on the right. Red lines represent the overlapped borders of the AAL regions at current sagittal, coronal and axial slices (39, 58, 39).

5.4.2   Evaluation of the number of features

We performed the classifications starting with the feature that had the highest ANOVA F-score and by adding the next highest scored feature in each step, are shown in 5.3. Ten features lead to the slightly better performance for both metrics.
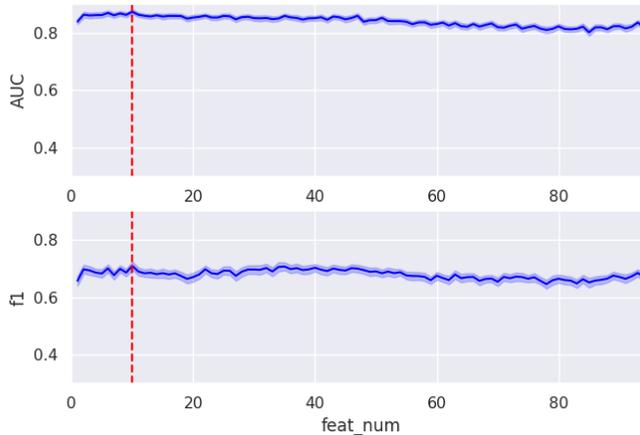
Figure 5.3: Optimal number of features was tested for the SVM classifier. We performed forward feature selection on the features ranked with the ANOVA F-scores. At first one features was included. For each step, the highest ranked feature not included was added, until all features are included. In every step AUC and F1-score was estimated. The combination of the 10 highest ranked features provided the best performance for both metrics.

5.4.3   Class weight evaluation

The desired 95% of *recall* which indicates 1 misclassified MCS patient out of twenty, corresponds to a cut off point where VS/UWS patients were weighted 2.4 times more than MCS. The 90% of *precision* occurs when VS/UWS where weighted 0.6 times of the MCS. We used those weight ratios to make two classifiers. The results across all selected weight-ratios are illustrated in figure 5.4.

5.4.4   Performance validation of selected models

The AUC for all classifiers was significantly higher than the chance level classification, as expressed by the *dummy classifier*. In all bootstrap iterations (1000) the AUC of the three classifiers was higher than the *dummy classifier*. Xtrees had lower *recall* compared to the *dummy* 4 times out of 1000. The *recall* of SVMp (sensitive in detecting VS/UWS) was lower than the *dummy*'s 344 times and SVMr's (sensitive in MCS detection), 1 time. Finally, for *precision* the performances of Xtrees and SVMp were higher than
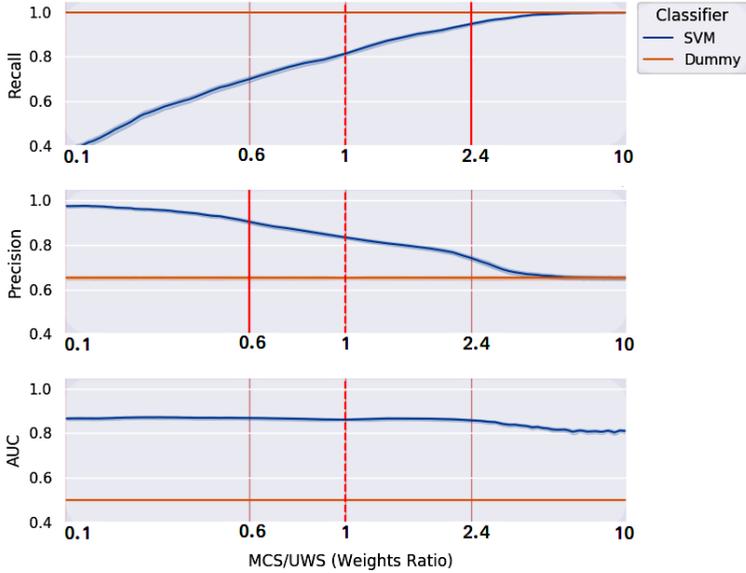
Figure 5.4: Estimating performance for different class weights in terms of *recall, Precision* and *AUC*. X axis represents the ratio of classification weights of MCS patients over those of VS/UWS. The red lines point out the weights at the 95% of *recall* and 90% of *precision*. A *dummy Classifier* was used to indicate the chance-level classification.

the *dummy*'s in all iterations, and SVMr performed 32 times worse than the *dummy classifier*. The mean performance for each metric and each classifier as well as the contrasts between each classifier and the *dummy* classifier are shown in figure 5.5.
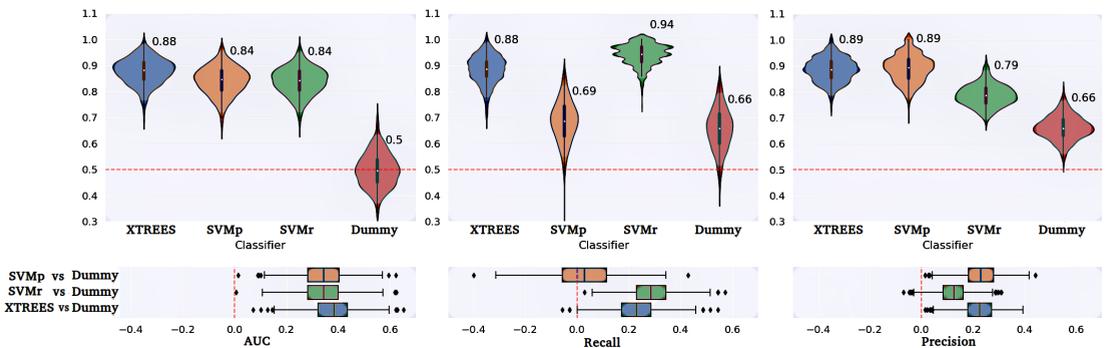
Figure 5.5: The performance validation of the selected models in terms of *recall, Precision* and *AUC*. Top three figures demonstrate the variance of the models and the *dummy classifier* for the three metrics. The three bottom boxplot-figures demonstrate the comparisons between the models and the *dummy classifier* for each iteration.

The table 5.1 summarizes the validation results and includes the 95% confidence intervals of the

bootstrap process.

Classifier	AUC (CI)	Recall (CI)	Precision (CI)
Xtrees	88% (77-96)	88% (77-97)	89% (78-97)
SVMp	84% (73-94)	69% (51-83)	89% (79-100)
SVMr	84% (73-94)	94% (86-1)	79% (71-88)
Dummy	50% (37-64)	66% (49-80)	66% (57-77)

Table 5.1: The performance on the validation set in terms of AUC, Recall and Precision. Metrics were estimated using bootstrap of 1000 iterations. In parenthesis the 95% confidence intervals for each metric are presented.

The performance of each model is shown in table 5.2. Models performed similarly to the performance achieved during the cross-validation scheme using the test-set.

		Models					
		SVM (recall)		SVM (precision)		Xtrees	
		MCS	UWS	MCS	UWS	MCS	UWS
Labels	MCS	33	2	24	11	31	4
	UWS	9	9	3	15	4	14
General accuracy		79%		74%		85%	
Recall		94%		69%		89%	
Precision		79%		89%		89%	

Table 5.2: Confusion matrices of models for the validation set. Classification results of the three selected models over the validation set, which included 35 MCS and 18 VS/UWS patients.

The table 5.3 provides the results of the final models which combines the three classifiers. Xtrees is used to provide a prediction only in the case that the two SVM models predict a different label for a subject. For the 38 out of 53 subjects, there was an agreement between the two SVM classifiers.

SVMr = SVMp		SVMr ≠ SVMp		
n = 38 (3FP, 2FN)		n = 15		
24 TP 2 FN	RF	7 TP 2 FN		
3 FP 9 TN		1 FP 5 TN		

Table 5.3: Confusion matrix of the combined classifiers. When the two SVM models agree 33/38 subjects are classified correctly. For the rest 15 cases, Xtrees classified correctly 12 out of 15 and had 1 FP and 2 FN. The total *recall* is 89%, *precision* is 89% and *Accuracy* is 85%

5.4.5 Contribution of regions in the classification

From the fitted models, we extracted the features significance and projected them on a 3-dimensional brain shaped model. The 10 regions that were selected using ANOVA F-scores used in SVM models are Paracentral Lobule, Cuneus, Lingual, Calcarine, Occipital Inferior, Occipital Superior, Precuneus, Postcentral in the right hemisphere and Cuneus with Calcarine on the left, and are shown in figure 5.6. Figure 5.7 shows the ranking of features (or regions) in the classification of the Xtrees. Xtrees ranks the features using the impurity decrement between nodes, averaged over all the trees.
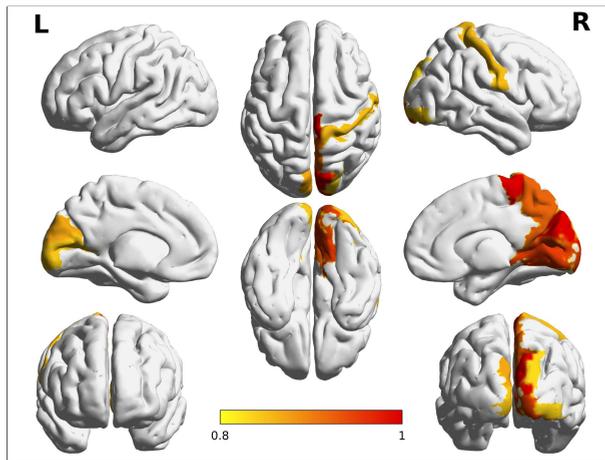
Figure 5.6: The 10 regions selected using ANOVA f-scores. F-scores values have been transformed with $-log_1 0$ and then normalized to unit for better visualization. Paracentral Lobule, Cuneus, Lingual, Calcarine, Occipital Inferior, Occipital Superior, Precuneus, Postcentral in the right hemisphere and Cuneus with Calcarine on the left. Regions in white (values below 0.82) were not selected in the best 10 features. Image was made with BrainNet viewer [169]

Figure 5.8 shows analytically the values of all regions for the two methods. The ANOVA F-scores have been transformed with $-log_{10}$ and normalized to one for better visualization. Xtrees have been only normalized to the unit.

5.4.6 Classification of healthy subjects

All 19 controls were classified as MCS from the Xtrees and the SVMr. One subject was classified as VS/UWS from the SVMp classifier.
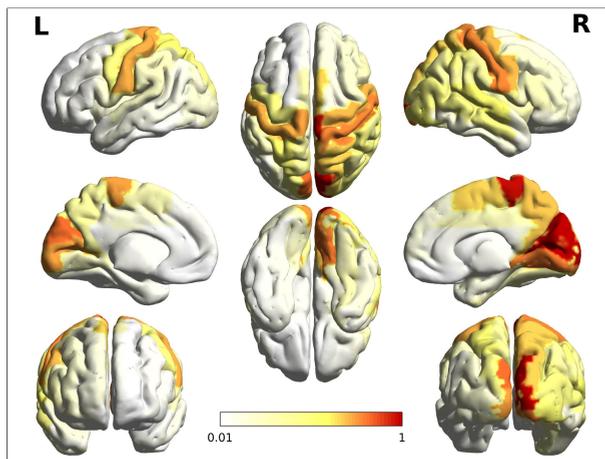
Figure 5.7: The importance of features as resulted from the Xtrees. The importance of features in Xtrees result from the nodes decrease of impurity averaged all over the trees. Mainly posterior and lateral brain areas drive the classifier. Similar regions from both sides of the brain are highlighted, with the right regions having higher values. Values have been normalized to 1. Image was made with BrainNet viewer [169]

5.5 Discussion

With this work we wanted to address the problem of an accurate diagnosis for DOC patients. Additionally, we wanted to investigate the regional contribution in the diagnosis of consciousness using a data driven approach. To overcome the problem of intensity variability between scans and permit comparison between subjects, scans were scaled to SUV. We also performed spatial normalization with a template created from FDG scans of subjects with brain lesions. We believe that such a template, compared to the standard SPM template, provides better warping of the images in the MNI space by preserving at the same time the lesions of the brain.

Specifications of the model

Though an optimal balance between *recall* and *precision* can be estimated with standardized metrics, in our case we opted to give a practical utilization. Therefore we set custom performance values to fit ethical and clinical expectations. By choosing the *recall* to be at 95% we actually defined the certainty level of classifying a patient as MCS to be correct in 19 out of 20 cases. A *precision* of 90% indicates that out of 10 detected MCS patients one will be VS/UWS. We assume that by penalizing differently the weights of one class over the other, the decision boundary would shift towards the direction of the class weighted less. Therefore, the two models will create almost parallel decision hyperplanes in the
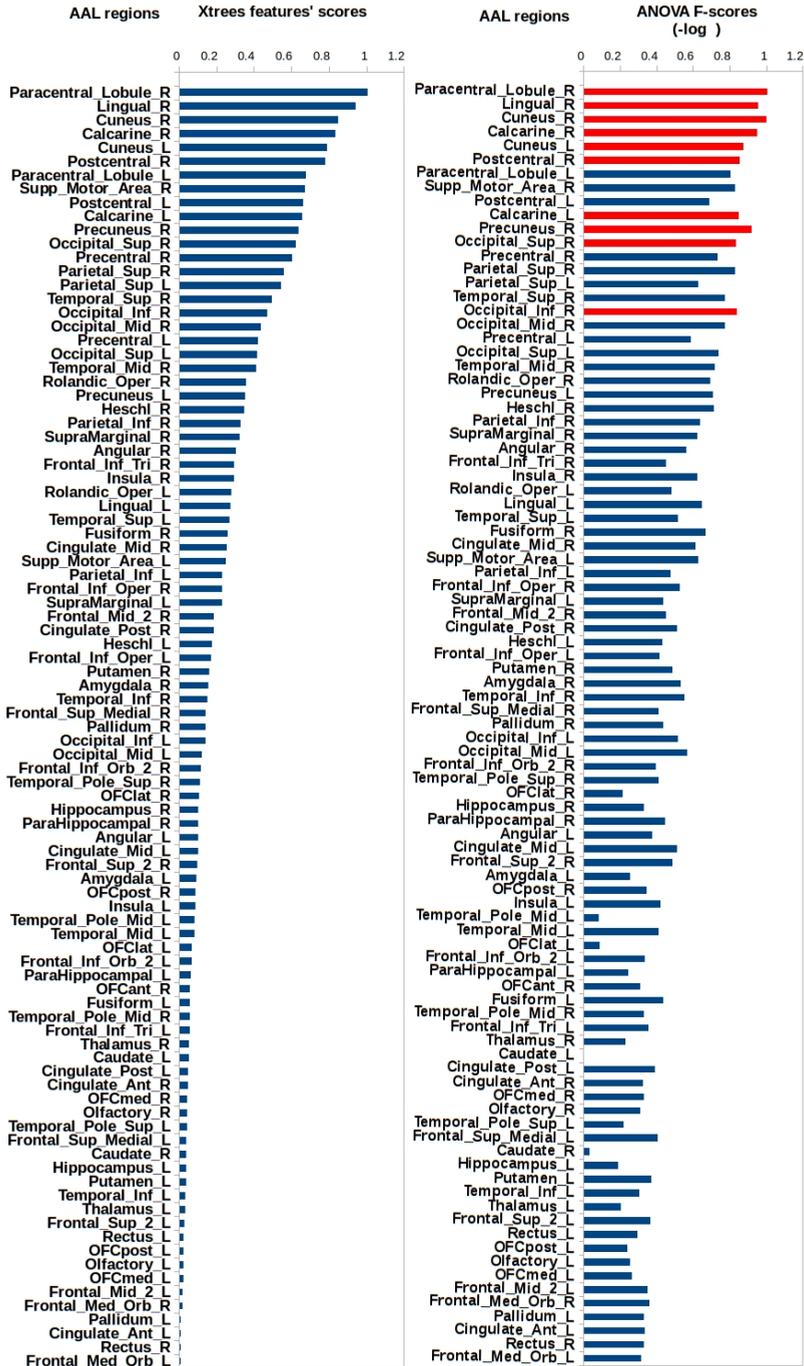
Figure 5.8: Values of features evaluation from the internal Xtrees classification process, on the right, and using ANOVA F-scores, on the left. In the ANOVA F-scores, the 10 features with the highest values are colored in red.

space of our interest. In such a way, when the two classifiers assign a sample with the same label, we can have a certain level of confidence about the classification result. When the two classifiers provide different labels, there are two possible ways for this to happen: 1) the high-recall SVM decides for MCS and the high-precision SVM for VS/UWS,and 2) the SVMr assigns the subject as VS/UWS and the SVMp as MCS. In the first case the subject would fall between the decision hyperplanes of the two classifiers, and will be decided in the Xtrees. The second scenario is unlikely and did not occur with any of our subjects, reaffirming in a way the assumption of having almost parallel hyperplanes. Figure 5.9 illustrates the aforementioned decision flow.
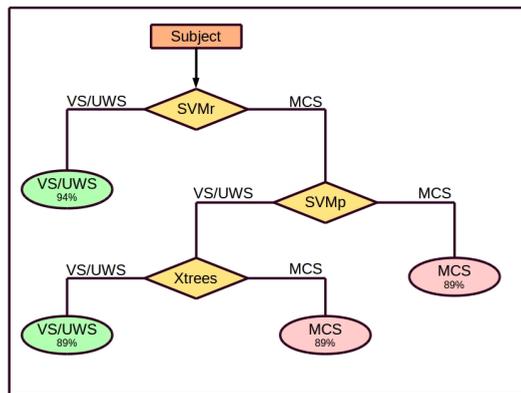
Figure 5.9: The flow chart shows the decision flow using all three classifiers. If the SVMr, which has low FN rate, assigns an VS/UWS, we know that is correct for 19 out of 20 cases. If the SVMr classifies the subject as MCS, we know that it's certainty level is low and have to proceed to the next SVMp classifier. The second classifier is sensitive in avoiding FP. It's certainty equals to one FP in every ten subjects classified as MCS. If this classifier decides also for MCS then this one can be the final decision. Finally, if the SVMp classifier decides for VS/UWS, then Xtrees will provide the final label.

It is of great importance that the selected performance was also achieved when using the validation dataset. That way we can be confident that the model performs as designed to perform on unseen data acquired with the same scanner and scanning protocol. The classification of controls as MCS was performed as a sanity check for the classifier. The one subject that was classified as VS/UWS showed very low mean compared to the rest of the controls which can be the result of sleeping between injection and scanning or protocol compliance 5.10.

The ANOVA F-scores analysis selected Paracentral Lobule, Cuneus, Lingual, Calcarine, Occipital Inferior, Occipital Superior, Precuneus, Postcentral in the right hemisphere and Cuneus with Calcarine
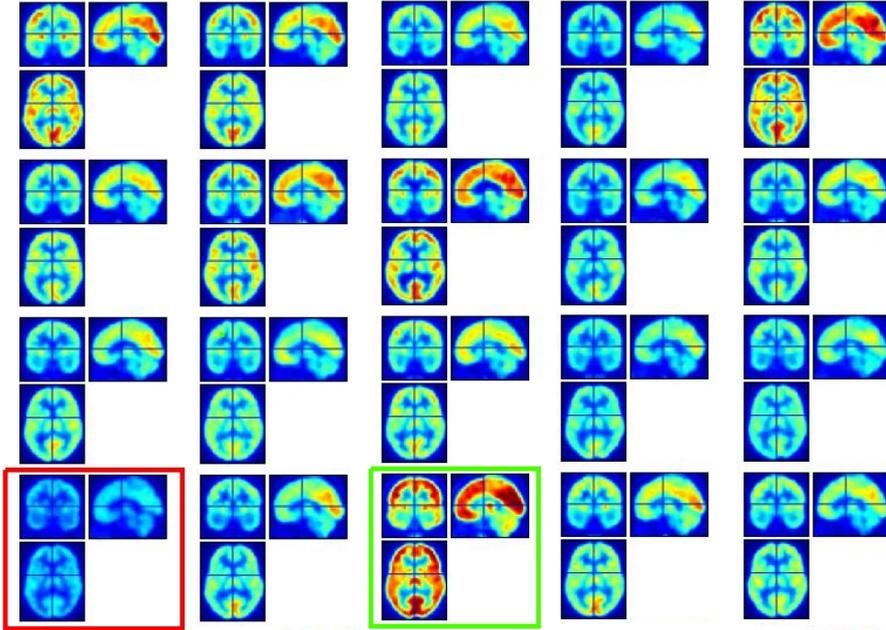
Figure 5.10: SUV scans of the healthy subjects. In the green frame is the subject that was excluded due to very high voxel values, which provided a z-score higher than 3. In the red frame is the subject that was classified as VS/UWS which has unexpectedly low values.

on the left. The regions selected are in accordance with the idea of "posterior hot zone" [170]. The Xtrees feature selection comes from averaging the decrease in impurity over trees. Faetures that ranked higher are Cuneus, Paracentral Lobule, Postcentral, Calcarine and the Precentral Parietal from both hemispheres, and Supplamentary Motor, Lingual, Precuneus and Occipital areas of the right hemisphere.

In both methods it can be noticed the significance of parietal and occipital midline, which have been highlighted in previous studies [51, 157] and right somatosensory cortex. For all classifiers the important regions for the classification are in both brain hemispheres although the right side is more highlighted.

The classification scheme we propose can compose the base for a more accurate model with fixed confidence intervals for every individual classification performed. Larger datasets with less noisy labels should increase the performance and moreover will permit for further development of the models. The area between the two hyperplanes is of great importance and a classifier with "Local Spacialization" [171], as part of multiple classifier scheme, can provide a more accurate overview of the differences of

the two groups.

In conclusion, we believe that the model proposed here can be used by clinicians and provide them with an assisting diagnostic tool. The classifier should always be used in combination with the clinical assessments and other evaluation methods if available.

Chapter 6

Discussion

6.1 General remarks

This work has shown how functional neuroimaging data -fMRI and PET- can be combined with machine learning in order to assist the diagnosis of the level of consciousness of brain lesioned patients. The first three introductory chapters provide fundamentals of the Disorders of Consciousness, functional neuroimaging focusing on PET and fMRI, and machine learning. The information provided are selected according to the techiques and tools used in the research described in the forth and fifth chapters. The forth chapter describes the performed analysis that resulted in a classification model which makes use of functional connectivity values. The fifth chapter includes the analysis I performed for diagnosing patients that are in MCS or VS/UWS using PET scans. To account for reproducability of the two models, the algorithms used have been described in detail and the corresponding code is available in version control system [1] or is provided upon request.

Although a big effort is made by the clinicians to provide an accurate diagnosis, there is no guarantee that all the patients involved are labeled correctly. This is a common problem in DOC research, which results in an upper limit of the classification performance. This limit exists, in my opinion, for all projects using clinical evaluations. Therefore, very high performances could imply overfitted models and validation with unseen data is in a way mandatory.

I believe that a more accurate labeling will come to fruition when clinical assessments will be used hand in hand with classification models and further tools coming from neuroimaging and neurophysiol-

---

[1] https://github.com/antogeo

ogy. The label refining for research purposes can be an added value for the two classification models presented in this work. With the time passing, bigger and cleaner databases will be created allowing the algorithms to provide a more accurate decision boundary, thereby increasing their performance as well as their robustness.

A significant parameter in the analysis of lesioned brains is the accurate preprocessing and segmentation. Most of the algorithms have been developed under the hypothesis that they will be applied in subjects with small or without distortions. Therefore, in DOC studies and especially for MRI, a significant number of subjects has to be excluded from the analysis. The PET images, are less affected by deformations due to their lower resolution and the fact that sedation can be applied without affecting the acquired data. Although the problem of big brain deformation has been addressed [162, 172] these solutions seem to be outdated. Additionally, it is not always possible to apply the proposed methodologies due to the time demanding manual intervention. Unfortunately, the inhomogeneities of the brains of DOC patients, caused by atrophies, lesions or traumas, do not permit us to transform the masks or the atlases in a way that they perfectly fit all subjects. I believe though, that potential mismatches between patients' brains and the masks or the parcellation atlases are small and moreover, are included in the variance that is taken into account by the models. There is a big debate whether or not the captured signal of BOLD in fMRI represents some local brain activity or is the result of an increased blood flow [173–175]. Also for PET, it is not clear whether the metabolic activity captured is purely coming from neurons or also from astrocytes [176, 177]. In both cases the validation of the classifiers was performed with unseen data. Therefore, I assume that the models encapsulated those factors in their parameters.

In the project of chapter 4 we tried to classify MCS from VS/UWS patients, using the functional connectivity of brain networks. Although the classifier differentiates conscious subjects from VS/UWS patients, it does not perform well with non-pathological lack of consciousness coming from anesthesia. Whether there are more ways to be unconscious, is a question that cannot be answered by our results, but raises an interesting aspect for further studies: is it possible to classify between unconscious subjects coming from pathological, pharmacological and physiological cause?

In PET classification the chosen scaling (SUV) is a standardized and widely used scaling technique despite the fact that in some case it can provide controversial results. The intensities of SUV-PET scans might have a wide range of values which is the result of inconsistency in data the registration and protocol compliance [178]. Relevance Uptake Value scaling, could overcome these issues and provide more accurate representation of the scans, but they need first to be studied and validated in subjects with big brain lesions and traumas. Accurate scaling can be achieved from methods that make use

of blood samples but they are very uncommon due to the high cost and difficulties in implementation. Hence, in our case SUV appeared to be the most reliable method.

Multivariate (Xtrees) and univariate (ANOVA F-scores) feature evaluation provide different but complementary information which facilitates discovering features which differently would be ignored [179, 180]. The former provides features importance and the latter shows if a variable is important either due to a correlation with the outcome or due to interdependencies with other variables that can be informative about the outcome [129]. Drawing conclusions about the weights assigned from the SVM can lead to mis-interpretetion. The regions highlighted in this project could constitute the basis of new rehabilitation research using targeted stimulation such as transcranial direct or alternate current stimulation.

6.2   Clinical Impact

The tools presented are using two different neuroimaging modalities and both can provide a diagnosis for the level of consciousness of patients to complement behavioural assessments. Due to the fact that PET and MRI have different scanning contraindications, the chance that one patient will not fulfill the scanning criteria for both of them is relatively small. Therefore, in most of the cases the clinical evaluation can be supported by one of the two classifiers.

In practice, if the behavioral assessments clearly indicate an MCS, there is no need to use the models. Clinicians mostly benefit from using the classifiers when other means of diagnosis indicate that a patient is unconscious. When both clinical evaluation and the models diagnose a patient as an VS/UWS the benefit of using the models is an increased confidence of the diagnosis. Therefore, patients care can be promoted by adjusting medical management like stimulating parts of the brain and decide about pain treatment. Additionally, it can help families to take decisions regarding end of life. More interesting though, can be those cases that one or both classifiers will not be in accordance with a clinical assessment indicating a VS/UWS, and classify a patient as an MCS. Then caregivers should consider further evaluations to be performed. Costly and complex set-ups such as "willful modulation" in an fMRI or other brain-computer interface methods should be planned and try to detect a response from the patient. This filtering of patients will lead to further examinations only for a targeted group of patients thus reducing the cost and improving the management of the resources.

6.3 Future work

For both projects future work includes the implementation of pipelines that will minimize manual intervention. The goal is that a user will have only to provide the DICOM images and a reorientation matrix as there is no algorithm so far that can provide an accurate reorientation of scans with big brain distortions. For the classifier of chapter 4, I would like to proceed to further tests with scans coming from meditating subjects, subjects under the influence of specific drugs or medication and examine if the way they are unconscious is similar to DOC patients. Additionally, I would like to search for a model that performs well with sedated patients which could be used for assisting the diagnosis of those patients.

For the model of chapter 5, I plan to test for the inter-scanner generalization as combining scanners is a big issue in PET studies [181]. Comparing different scaling methods, SUV and Relevance Uptake Value (RUV), and testing them in terms of classification performance is also a future perspective. An accurate RUV scaling will set unnecessary all the information needed for the SUV scaling. Finally, I intend to combine classifiers and create ensembles of classifiers of different modalities. Combining functional and structural information has already shown that increases performance [182]. The main idea consists of a multimodal model using the two classifiers of this work, a classifier using structural information [183] and one using neurophysiology [115]. In such multiple classification system classifiers can have a parallel structure and vote for the final decision, or serial structure where one is feeding the next one with some *a priori* information. The weights of each vote or the sequence of the classifiers result from test.

Chapter 7

Paper I

Intrinsic functional connectivity differentiates minimally conscious from unresponsive patients

Demertzi A*, Antonopoulos G*, Heine L, Voss H U, Crone S J, Kronbichler M, Trinka E, Angeles C, Bahri M, Phillips C, Di-Perri C, Gomez F, Tshibanda L, Soddu A, Vanhaudenhuyse A, Charland-Verville V, Schiff N D, Whitfield-Gabrieli S*, Laureys S*.

BRAIN
A JOURNAL OF NEUROLOGY

Intrinsic functional connectivity differentiates minimally conscious from unresponsive patients

Athena Demertzi,[1,]* Georgios Antonopoulos,[1,]* Lizette Heine,[1] Henning U. Voss,[2]
Julia Sophia Crone,[3,4,5] Carlo de Los Angeles,[6] Mohamed Ali Bahri,[7] Carol Di Perri,[1]
Audrey Vanhaudenhuyse,[8] Vanessa Charland-Verville,[1] Martin Kronbichler,[3,4]
Eugen Trinka,[5] Christophe Phillips,[7] Francisco Gomez,[9] Luaba Tshibanda,[10]
Andrea Soddu,[11] Nicholas D Schiff,[12,13] Susan Whitfield-Gabrieli[6,]* and Steven Laureys[1,]*

*These authors contributed equally to this work.

Despite advances in resting state functional magnetic resonance imaging investigations, clinicians remain with the challenge of how to implement this paradigm on an individualized basis. Here, we assessed the clinical relevance of resting state functional magnetic resonance imaging acquisitions in patients with disorders of consciousness by means of a systems-level approach. Three clinical centres collected data from 73 patients in minimally conscious state, vegetative state/unresponsive wakefulness syndrome and coma. The main analysis was performed on the data set coming from one centre (Liège) including 51 patients (26 minimally conscious state, 19 vegetative state/unresponsive wakefulness syndrome, six coma; 15 females; mean age 49 ± 18 years, range 11–87; 16 traumatic, 32 non-traumatic of which 13 anoxic, three mixed; 35 patients assessed >1 month post-insult) for whom the clinical diagnosis with the Coma Recovery Scale-Revised was congruent with positron emission tomography scanning. Group-level functional connectivity was investigated for the default mode, frontoparietal, salience, auditory, sensorimotor and visual networks using a multiple-seed correlation approach. Between-group inferential statistics and machine learning were used to identify each network's capacity to discriminate between patients in minimally conscious state and vegetative state/unresponsive wakefulness syndrome. Data collected from 22 patients scanned in two other centres (Salzburg: 10 minimally conscious state, five vegetative state/unresponsive wakefulness syndrome; New York: five minimally conscious state, one vegetative state/unresponsive wakefulness syndrome, one emerged from minimally conscious state) were used to validate the classification with the selected features. Coma Recovery Scale-Revised total scores correlated with key regions of each network reflecting their involvement in consciousness-related processes. All networks had a high discriminative capacity (>80%) for separating patients in a minimally conscious state and vegetative state/unresponsive wakefulness syndrome. Among them, the auditory network was ranked the most highly. The regions of the auditory network which were more functionally connected in patients in minimally conscious state compared to vegetative state/unresponsive wakefulness syndrome encompassed bilateral auditory and visual cortices. Connectivity values in these three regions discriminated congruently 20 of 22 independently assessed patients. Our findings point to the significance of preserved abilities for multisensory integration and top–down processing in minimal consciousness seemingly supported by auditory-visual crossmodal connectivity, and promote the clinical utility of the resting paradigm for single-patient diagnostics.

1   Coma Science Group, GIGA-Research & Cyclotron Research Centre, University and CHU University Hospital of Liège, Liège, Belgium
2   Department of Radiology and Citigroup Biomedical Imaging Centre, Weill Cornell Medical College, New York, USA
3   Department of Psychology and Centre for Neurocognitive Research, Salzburg, Austria

4 Neuroscience Institute and Centre for Neurocognitive Research, Christian-Doppler-Klinik, Paracelsus Private Medical University, Salzburg, Austria
5 Department of Neurology, Christian-Doppler-Klinik, Paracelsus Private Medical University, Salzburg, Austria
6 Martinos Imaging Centre at McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge MA, USA
7 Cyclotron Research Centre, University of Liège, Liège, Belgium
8 Department of Algology and Palliative Care, CHU University Hospital of Liège, Liège, Belgium
9 Computer Science Department, Universidad Central de Colombia, Bogota, Colombia
10 Department of Radiology, CHU University Hospital of Liège, Liège, Belgium
11 Brain and Mind Institute, Department of Physics and Astronomy, Western University, London, Ontario, Canada
12 Department of Neuroscience, Weill Cornell Graduate School of Medical Sciences, New York, USA
13 Department of Neurology and Neuroscience, Weill Cornell Medical College, New York, USA

Correspondence to: Demertzi Athena,
Coma Science Group, Cyclotron Research Centre,
Allée du 6 août n° 8, Sart Tilman B30,
Université de Liège, 4000 Liège, Belgium.
E-mail: a.demertzi@ulg.ac.be

Introduction

As patients with acute or chronic disorders of consciousness are by definition unable to communicate, their diagnosis is particularly challenging. Patients in coma, for example, lay with eyes closed and do not respond to any external stimulation. When they open their eyes but remain unresponsive to external stimuli they are considered to be in a vegetative state (VS; Jennett and Plum, 1972) or, as most recently coined, unresponsive wakefulness syndrome (UWS; Laureys *et al.*, 2010). When patients exhibit signs of fluctuating yet reproducible remnants of non-reflex behaviour, they are considered to be in a minimally conscious state (MCS; Giacino *et al.*, 2002). To date, the diagnostic assessment of patients with disorders of consciousness is mainly based on the observation of motor and oro-motor behaviours at the bedside (Giacino *et al.*, 2014). The evaluation of non-reflex behaviour, however, is not straightforward as patients can fluctuate in terms of vigilance, may suffer from cognitive (e.g. aphasia, apraxia) and/or sensory impairments (e.g. blindness, deafness), from small or easily exhausted motor activity and pain. In these cases, absence of responsiveness does not necessarily correspond to absence of awareness (Sanders *et al.*, 2012). Alternatively, motor-independent technologies can aid the clinical differentiation between the two patient groups (Bruno *et al.*, 2010).

Up to now, accurate single-patient categorization in MCS and VS/UWS has been performed by means of transcranial magnetic stimulation in combination with EEG (Rosanova *et al.*, 2012; Casali *et al.*, 2013) and by combining different EEG measures (Sitt *et al.*, 2014). In terms of patient separation by means of functional MRI, activation (which utilise sensory stimulation; Schiff *et al.*, 2005; Coleman *et al.*, 2007; Di *et al.*, 2007) and active paradigms (which probe

mental command following; Owen *et al.*, 2006; Monti *et al.*, 2010; Bardin *et al.*, 2012) have been used to detect convert awareness in these patients. An apparent limitation of the latter approaches is that patients may demonstrate motor and language deficits which incommode these assessments and heighten the risk of false-negative findings (Giacino *et al.*, 2014). The application of these paradigms can also be constrained due to each institution's technical facilities.

Alternatively, functional MRI acquisitions during resting state do not require sophisticated setup and surpass the need for subjects' active participation. Past resting state functional MRI-based assessment of patients has focused on the default mode network, which mainly encompasses anterior and posterior midline regions, and which has been involved in conscious and self-related cognitive processes (Raichle *et al.*, 2001; Buckner *et al.*, 2008). Such investigations have shown that default mode network functional connectivity decreases alongside the spectrum of consciousness, moving from healthy controls to patients in MCS, VS/UWS and coma (Boly *et al.*, 2009; Vanhaudenhuyse *et al.*, 2010; Norton *et al.*, 2012; Soddu *et al.*, 2012; Demertzi *et al.*, 2014; Huang *et al.*, 2014). In patients, the precuneus and posterior cingulate cortex of the default mode network have been also characterized by decreases in functional MRI resting state low frequency fluctuations and regional voxel homogeneity (which refers to the similarity of local brain activity across a region) (Tsai *et al.*, 2014). Reduced functional MRI functional connectivity has been further identified for interhemispheric homologous regions belonging to the extrinsic or task-positive network (implicated in the awareness of the environment; Vanhaudenhuyse *et al.*, 2011) in patients as compared to controls (Ovadia-Caro *et al.*, 2012). Reduced interhemispheric connectivity has

been also indicated by means of partial correlations (Maki-Marttunen *et al.*, 2013). In terms of graph theory metrics, comatose patients were shown to preserve global network properties but cortical regions, which worked as hubs in healthy controls, became non-hubs in comatose brains and vice versa (Achard *et al.*, 2011, 2012). Similarly, chronic patients showed altered network properties in medial parietal and frontal regions as well as in the thalamus, and most of the affected regions in unresponsive patients belonged to the so-called 'rich-club' of highly interconnected central nodes (Crone *et al.*, 2014). More recently, functional MRI-based single-patient classification has been performed by considering as discriminating feature the neuronal properties of various intrinsic connectivity networks (Demertzi *et al.*, 2014). The discrimination between 'neuronal' and 'non-neuronal' was based on the spatial and temporal properties (fingerprints) of the identified networks that were extracted by means of independent component analysis (De Martino *et al.*, 2007). According to specific criteria (Kelly *et al.*, 2010), 'non-neuronal' components were those that showed activation/deactivation in peripheral brain areas, in the cerebrospinal fluid (CSF) and white matter, as well as those showing high frequency fluctuations (>0.1 Hz), spikes, presence of a sawtooth pattern and presence of thresholded voxels in the superior sagittal sinus. Conversely, 'neuronal' were those networks when at least 10% of the activations/deactivations were found in small to larger grey matter clusters localized to small regions of the brain. Based on this definition of neuronality, the 'neuronal' properties of the default mode and auditory network were able to separate single-patients from healthy controls with 85.3% accuracy. Nevertheless, the discrimination accuracy between patients in MCS and VS/UWS reached only a chance level (Demertzi *et al.*, 2014).

Taken together, these studies show that the so far resting state functional MRI-based differentiation of patients has been performed either at the group-level or concerned the classification between healthy and pathological groups. As a consequence, clinicians remain with the challenge of how to implement the resting state functional MRI paradigm on an individualized basis for the more challenging discrimination between the MCS and VS/UWS (Edlow *et al.*, 2013). Here, we aimed at promoting the MCS-VS/UWS single-patient differentiation by using resting state functional MRI measurements in this clinical population. To this end, we studied systems-level resting state functional MRI functional connectivity in traumatic and non-traumatic patients with acute and chronic disorders of consciousness with the aim to (i) estimate the contribution of each network to the level of consciousness as determined by behavioural assessment; (ii) rank the capacity of each network to differentiate between patients in MCS and VS/UWS; and (iii) automatically classify independently assessed patients.

Materials and methods

Subjects

Three data sets were used, including patients scanned in Liège [to address study aims (i) and (ii)], Salzburg and New York [to address study aim (iii)]. Inclusion criteria were patients in MCS, VS/UWS and coma following severe brain damage studied at least 2 days after the acute brain insult. Patients were excluded when there was contraindication for MRI (e.g. presence of ferromagnetic aneurysm clips, pacemakers), MRI acquisition under sedation or anaesthesia, and uncertain clinical diagnosis. Healthy volunteers were free of psychiatric or neurological history. The study was approved by the Ethics Committee of the Medical School of the University of Liège, the Ethics Committee of Salzburg, and the Institutional Review Board at Weill Cornell Medical College. Informed consent to participate in the study was obtained from the healthy subjects and from the legal surrogates of the patients.

Data acquisition

All data were acquired on 3 T Siemens TIM Trio MRI scanners (Siemens Medical Solutions). For the Liège data set, 300 multislice T_2^*-weighted images were acquired with a gradient-echo echo-planar imaging sequence using axial slice orientation and covering the whole brain (32 slices; voxel size = $3 \times 3 \times 3$ mm³; matrix size = 64×64; repetition time = 2000 ms; echo time = 30 ms; flip angle = 78°; field of view = 192×192 mm). For the Salzburg data set, 250 T_2^*-weighted images (36 slices with 3-mm thickness; repetition time = 2250 ms; echo time = 30 ms; flip angle = 70°; field of view = 192×192 mm). For the New York data set, 180 T_2^*-weighted images were acquired (32 slices; voxel size = $3.75 \times 3.75 \times 4$ mm³; matrix size = 64×64; repetition time = 2000 ms; echo time = 30 ms; flip angle = 90°; field of view = 240×240 mm).

Subject-level connectivity analysis

Data analysis is illustrated in Fig. 1.

Data preprocessing

Preprocessing and connectivity analyses were performed in the same way for all subjects across the three data sets. The three initial volumes were discarded to avoid T_1 saturation effects. For anatomical reference, a high-resolution T_1-weighted image was acquired for each subject (T_1-weighted 3D magnetization-prepared rapid gradient echo sequence). Data preprocessing was performed using Statistical Parametric Mapping 8 (SPM8; www.fil.ion.ucl.ac.uk/spm). Preprocessing steps included slice-time correction, realignment, segmentation of structural data, normalization into standard stereotactic Montreal Neurological Institute (MNI) space and spatial smoothing using a Gaussian kernel of 6 mm full-width at half-maximum. As functional connectivity is influenced by head motion in the scanner (Van Dijk *et al.*, 2012), we accounted for motion artifact detection and rejection using the artifact detection tool (ART; http://www.nitrc.org/projects/artifact_detect). Specifically, an image was defined as an outlier (artifact) image if the head displacement in x, y, or

Figure 1 Analysis pipeline. Data analysis at the subject-level encompassed signal preprocessing and extraction of the intrinsic connectivity networks. Data analysis at the group-level encompassed estimation of functional connectivity in the networks of interest, estimated the contribution of each network to the level of consciousness by means of CRS-R total score regression analysis, and identified connectivity differences between the group of patients in MCS and VS/UWS for each network. Network ranking methodology was used to rank characteristic features (i.e., connectivity differences per network) to discriminate individual patients into the groups of MCS and VS/UWS. Two independent data sets of patients, assessed in Salzburg and New York, were used to further validate patient classification. Different colours indicate the three data sets and how these where used along the analysis pipeline.

z direction was >0.5 mm from the previous frame, or if the rotational displacement was >0.02 radians from the previous frame, or if the global mean intensity in the image was >3 standard deviations (SD) from the mean image intensity for the entire resting scan. Outliers in the global mean signal intensity and motion were subsequently included as nuisance regressors (i.e. one regressor per outlier within the first-level general linear model). Therefore, the temporal structure of the data was not disrupted.

For noise reduction, previous methods subtracted the global signal across the brain (a controversial issue in resting state analyses; Murphy *et al.*, 2009; Saad *et al.*, 2012; Wong *et al.*, 2012), and the mean signals from noise regions of interest (Greicius *et al.*, 2003; Fox *et al.*, 2005). Here, we used the anatomical component-based noise correction method (aCompCor; Behzadi *et al.*, 2007) as implemented in CONN functional connectivity toolbox (http://www.nitrc.org/projects/conn/; Whitfield-Gabrieli and Nieto-Castanon, 2012). The aCompCor models the influence of noise as a voxel-specific linear combination of multiple empirically estimated noise sources by deriving principal components from noise regions of interest and by including them as nuisance parameters within the general linear models. Specifically, the anatomical image for each participant was segmented into white matter, grey matter, and CSF masks using SPM8. To minimize partial voluming with grey matter, the white matter and CSF masks were eroded by one voxel, which resulted in substantially smaller masks than the original segmentations (Chai *et al.*, 2012). The eroded white matter and CSF masks were then

used as noise regions of interest. Signals from the white matter and CSF noise regions of interest were extracted from the unsmoothed functional volumes to avoid additional risk of contaminating white matter and CSF signals with grey matter signals. A temporal band-pass filter of 0.008–0.09 Hz was applied on the time series to restrict the analysis to low frequency fluctuations, which characterize functional MRI blood oxygenation level-dependent resting state activity as classically performed in seed-correlation analysis (Greicius *et al.*, 2003; Fox *et al.*, 2005). Residual head motion parameters (three rotation and three translation parameters, plus another six parameters representing their first-order temporal derivatives) were regressed out.

Extraction of intrinsic connectivity networks

Functional connectivity adopted a seed-based correlation approach. Seed-correlation analysis uses extracted blood oxygenation level-dependent time series from a region of interest (the seed) and determines the temporal correlation between this signal and the time series from all other brain voxels. Evidently, the selection of the seed region is critical because, in principle, it can lead to as many overlapping networks as the number of possible selected seeds (Cole *et al.*, 2010). Additionally, a network disruption can be expected due to patients' underlying neuropathology, as the chosen seed may no longer be included in the overall network. Using more seed regions, this issue can be overcome and therefore ensure proper network characterization in patients. Here, the seeds

that were selected to replicate the networks were defined as 10-mm (for cortical areas) and 4-mm radius spheres (for subcortical structures) around peak coordinates taken from the literature (Supplementary material). For each network, time series from the voxels contained in each seed region were extracted and then averaged together. In that way, the resulting averaged time course was estimated by taking into account the time courses of more than one regions. The averaged time series were used to estimate whole-brain correlation r maps that were then converted to normally distributed Fisher's z transformed correlation maps to allow for group-level comparisons.

Group-level connectivity analysis

For the Liège data set, one-sample *t*-tests were ordered to estimate network-level functional connectivity for patients in MCS, VS/UWS and in coma; the data from healthy controls were used as a reference to ensure proper network characterization. An exploratory analysis looked for network-level connectivity changes as a function of patients' aetiology and chronicity. Two 2×2 factorial designs between aetiology (traumatic, non-traumatic)/ chronicity (acute, chronic) and the clinical entities (MCS, VS/UWS) were ordered. If an interaction effect was identified, these variables had to be entered as regressors in the general linear models.

To address the first aim of the study, i.e. to estimate the contribution of each network to the level of consciousness, patients' Coma Recovery Scale-Revised (CRS-R) total scores were used as regressors to determine the relationship between each network's functional connectivity and the level of consciousness. As a control, CRS-R total scores were used as regressors of functional connectivity for the cerebellum network (three regions of interest, Supplementary material), which is known to be minimally implicated in consciousness-related processes (Tononi, 2008; Yu *et al.*, 2015).

To address the second aim of the study, i.e. to determine the capacity of each network to differentiate between patients in MCS and VS/UWS, initially two-sample *t*-tests were ordered to identify the regions of each network showing higher functional connectivity in patients in MCS compared to VS/UWS (Liège data set). The resulting difference maps were saved as masks, which were used subsequently for the network ranking and selection step. All results were considered significant $P < 0.05$ corrected for multiple comparisons at false discovery rate (FWE; cluster-level).

Network ranking and selection

Using the REX Toolbox (http://www.nitrc.org/projects/rex/), the difference masks which were calculated in the previous step were used to extract mean connectivity values (average z-values across the whole mask) from the first-level contrast images estimated for each network. Therefore, one value per subject per network was created leading to a 6×1 vector per subject (i.e. 45×6 matrix). These vector values were considered as features in a feature ranking methodology (Saeys *et al.*, 2007) as implemented in Matlab (http://www.mathworks.nl/help/bioinfo/ref/rankfeatures.html). The results of the feature (i.e. network) ranking were verified by means of single-feature linear support vector machine classifier (Burges, 1998). Supplementary material contains further details on the network ranking procedure and results.

To address the third aim of the study, i.e. to automatically classify independently assessed patients coming from two other clinical centres, we focused on the network which was ranked most highly during the network ranking procedure. For that network, a linear kernel support vector machine classifier (Burges, 1998) with regularization parameter C = 1 was used. This parameter was chosen based on its wide use in the machine learning procedure (Phillips *et al.*, 2011). The features that were used for the training were individual mean connectivity values extracted from the first-level contrast images using the relevant network binary mask as described above. To avoid single feature classification, hence running the risk of overfitting, more features were included for the classifier's training. The number of features was based on the number of clusters showing higher connectivity in patients in MCS compared to VS/UWS as indicated by the contrast manager of the CONN toolbox during the connectivity analysis (FWE $P < 0.05$, cluster-level correction).

Classification of independently assessed patients

The final validation of the classifier was performed on a new set of connectivity values extracted from independently assessed patients in Salzburg ($n = 15$) and New York ($n = 7$). The data preprocessing, extraction of intrinsic connectivity network, and feature extraction followed an identical procedure as described above for the Liège data set. To test for robustness, we also evaluated whether the same classifier generalized to healthy controls subjects scanned in two centres (Liège, Salzburg; no healthy control data were available for the New York centre).

Results

Subjects

In Liège, between April 2008 and December 2012, 177 patients with disorders of consciousness underwent MRI scanning. Of these, 80 (45%) were excluded due to sedation or anaesthesia during scanning. Of the remaining 97 patients scanned in an awake state, five due to change of diagnosis within a week after scanning, 14 because they showed functional communication, 15 due to technical reasons or movement artifacts, and 12 due to incongruence between clinical diagnosis and fluorodeoxyglucose (FDG)-PET scanning (Stender *et al.*, 2014). As regards the latter criterion, we decided to exclude patients showing widespread PET activation in midline and frontoparietal regions while the bedside diagnosis indicated the VS/UWS, in order to avoid confounds due to clinical ambiguity.

The included 51 patients were behaviourally diagnosed with the CRS-R (Giacino *et al.*, 2004) as in MCS = 26, VS/UWS = 19 and coma = 6 (15 females; mean age 49 ± 18 years, range 11–87; 16 traumatic, 32 non-traumatic of which 13 were anoxic, three mixed; 35 patients were assessed in the chronic setting, i.e. >1 month post-insult). Data from an age-matched group of 21 healthy volunteers

(eight females; mean age 45 ± 17 years; range 19–72) were used as a reference to the connectivity analyses and to validate the generalizability of the classifier without being included in the training. The data set from Salzburg included 10 MCS and five VS/UWS patients; the data set from New York included five MCS, one VS/UWS and one patient emerged from MCS. All patients' demographic and clinical characteristics are summarized in the Supplementary material.

For the Liège data set, the effects of the denoising procedure are summarized in the Supplementary material. Also, the number of motion outlier images did not differ among healthy controls (mean = 9 ± 8), patients in MCS (mean = 22 ± 17), VS/UWS (mean = 17 ± 12), coma (mean = 2 ± 2) (for all *t*-tests, *P* < 0.05). The exploratory analysis indicated a main effect for the clinical entity (i.e. MCS, VS/UWS) on the functional connectivity of each network. No interaction was identified between the clinical entity and aetiology (traumatic: MCS = 13, VS/UWS = 1; non-traumatic: MCS = 12 + 1 mixed; VS/UWS = 16 + 2 mixed) or chronicity (acute MCS = 5, VS/UWS = 6; chronic MCS = 21, VS/UWS = 13; average length of time since the injury was 902.3 days, minimum = 2 days, maximum = 9900).

Group-level connectivity analysis

For the default mode, frontoparietal, salience, auditory, sensorimotor and visual network, functional connectivity encompassed regions classically reported for healthy controls; all six networks showed reduced connectivity in patients in MCS, connectivity was hardly identified in patients in VS/UWS and was absent in comatose patients (Supplementary material).

CRS-R total scores correlated with functional connectivity in key regions of each network (Fig. 2). In contrast, when the CRS-R total scores were used as regressors of connectivity in the cerebellum, which is known for its minimal involvement in consciousness processes (Tononi, 2008), no areas showed connectivity with the behavioural scores. For illustrative purposes, the cerebellar network in healthy controls is presented in the Supplementary material.

The regions that showed higher functional connectivity in patients in MCS compared to VS/UWS for each network are summarized in Fig. 3. To minimize the possibility that differences in functional connectivity reflected differences in brain anatomy, we performed a two-sample *t*-test voxel-based morphometry on the normalized grey matter and white matter segmented masks (smoothed at 6 mm full-width at half-maximum). No differences in grey matter volume between patients in MCS and VS/UWS were identified at FWE *P* < 0.05 either at the whole-brain or at the cluster-level. Similarly, the analysis of white matter volumes identified no differences between the two groups, even at a liberal threshold *P* < 0.001 (whole brain level) uncorrected for multiple comparisons. The average grey matter and white matter volumes in the two patient groups are reported in the Supplementary material.

Network ranking and selection

All networks were found to discriminate between patients in MCS and VS/UWS with an acceptable accuracy (Supplementary material). Among them, the auditory network was the most highly ranked system to separate patients in MCS from those in VS/UWS.

## Validation with independent data set

Functional connectivity of the auditory network was further used to classify independently assessed patients. The classification was performed on the connectivity strength in bilateral auditory and visual cortices (Fig. 3). This three-feature vector was preferred to a single-feature classification (i.e. the average connectivity across all areas of the auditory network mask) to avoid over-fitting of the classifier. Based on these three clusters' connectivity strength (*z*-values), 20 of 22 patients independently assessed in Salzburg and New York were discriminated congruently (Fig. 4 and Supplementary material), namely the CRS-R diagnosis matched the classification outcome. As in Phillips *et al.* (2011), for each feature we calculated its weighted vector 'w', which determines the orientation of the decision surface, indicative of which feature drives the classification (Bishop, 2006). For the right auditory cortex it was w = −1.7890, for the left auditory cortex w = −0.4002 and for the occipital cortex w = −0.7362. The patient who was misclassified as being in MCS had a CRS-R total score of 5 on the day of scanning (indicating the VS/UWS; Patient 11 of centre two, Supplementary material) and she evolved to MCS 38 days later (Auditory Function: 1, Visual Function: 3, Motor Function: 2, Oromotor/Verbal Function: 2, Communication: 0, Arousal: 2). The patient who was misclassified as being in VS/UWS had a CRS-R total score of 9 on the day of scanning (indicating the MCS; Patient 13 of centre two, Supplementary material) based on the presence of localization to noxious stimulation but this behaviour could not be elicited in neither previous (AF: 1, VF: 0, MF: 0, O/VF: 1, COM: 0, AR: 2) or subsequent evaluations (AF: 2, VF: 1, MF: 2, O/VF: 1, COM: 0, AR: 2). To test robustness, we evaluated whether the same classifier generalized to healthy control subjects scanned in Liège and Salzburg (*n* = 39; no healthy control data were available for the New York centre). The majority of healthy controls (37 of 39; 95%) were classified as MCS (Supplementary material).

Discussion

We here aimed at determining the clinical utility of the resting state functional MRI paradigm in patients with disorders of consciousness by employing a systems-level
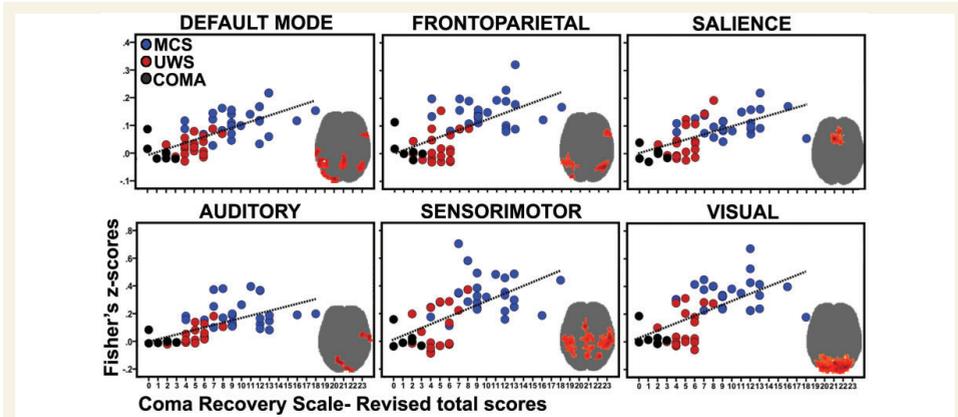
Figure 2 The intrinsic connectivity networks are involved in consciousness-related processing. Functional connectivity of all studied networks (areas in red) correlate with the level of consciousness as determined by behavioural assessment with the Coma Recovery Scale-Revised (total scores) in patients in MCS, VS/UWS and coma. Statistical maps are thresholded at FWE $P < 0.05$ (cluster-level) and are rendere on a glass brain template (transverse view).
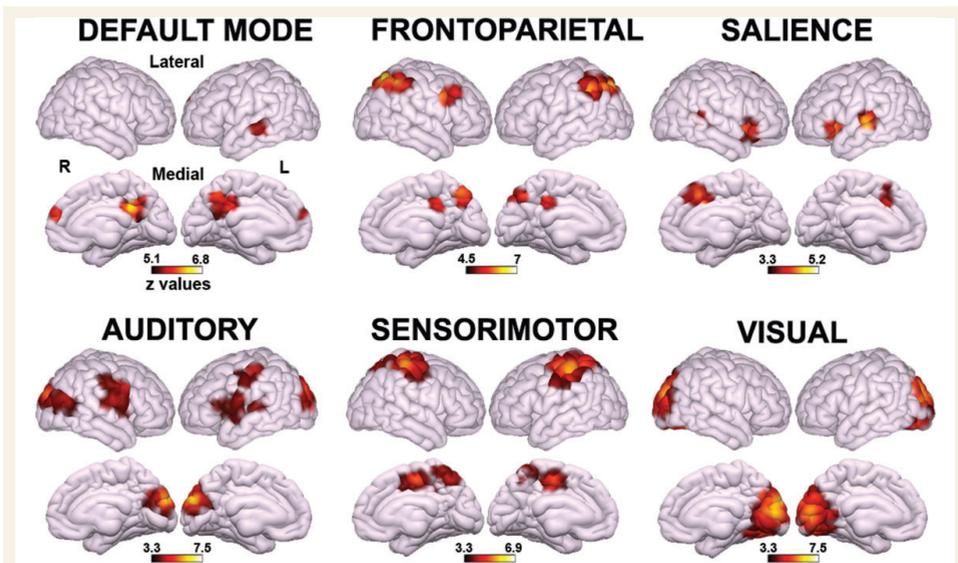
Figure 3 Regions showing higher functional connectivity in patients in MCS compared to patients in VS/UWS for each network. Statistical maps are thresholded at FWE $P < 0.05$ (cluster-level) and are rendered on 3D surface plot template (*top* = lateral view; *bottom* = medial view).

approach. Resting state functional MRI connectivity of the default mode, frontoparietal, salience, auditory, sensorimotor and visual networks were first shown to correlate with behavioural CRS-R assessment scores, highlighting

their contribution to the level of consciousness. Previous studies on the default mode network, linked to autobiographical memory, mind-wandering, and unconstrained cognition (Buckner *et al.*, 2008), also showed
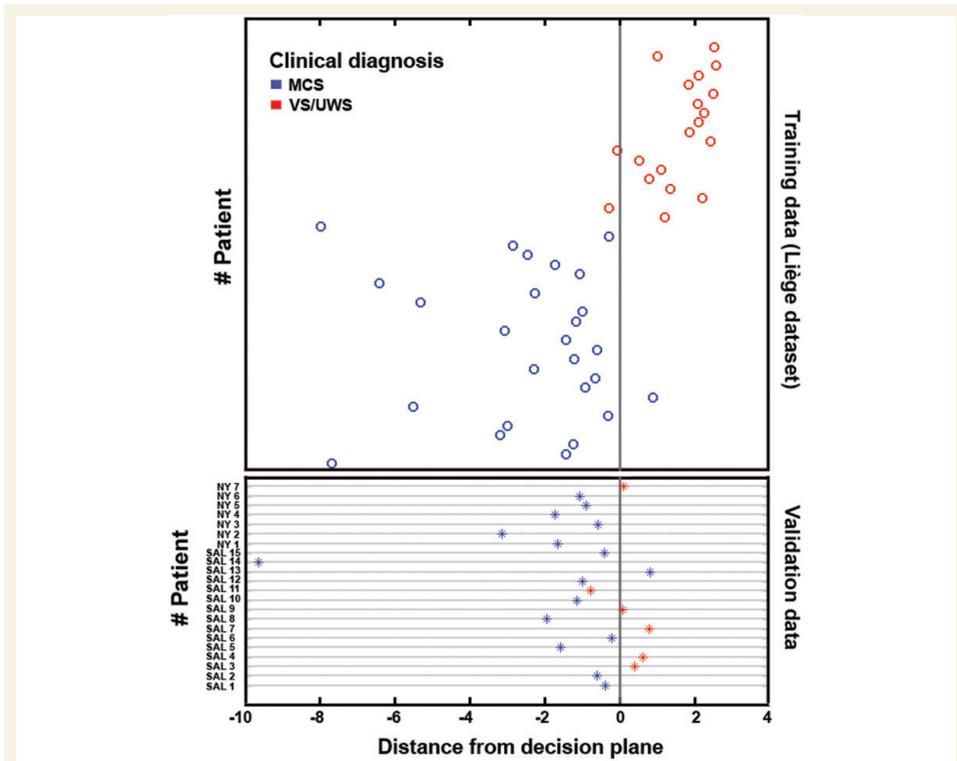
Figure 4 The auditory-visual crossmodal functional connectivity discriminates single patients in MCS from patients in VS/UWS. The 3D space indicating connectivity between left auditory, right auditory and occipital cortex (Supplementary material) has been compressed into two dimensions to represent the distance of each patient (in circles) from the decision plane (arbitrary values). The *upper* panel plots the data of patients (in circles) who were used for the classifier's training (Liège data set, *n* = 45). The *lower* panel summarizes the classifier's decision on the validation data set including patients (in asterisks) independently assessed in Salzburg (*n* = 15) and New York (*n* = 7). Based on the crossmodal interaction, 20 of the 22 independently assessed patients were classified congruently, namely the behavioural diagnosis matched the classification outcome.

consciousness-level dependent reductions in connectivity under physiological (Horovitz *et al.*, 2009; Samann *et al.*, 2011) and pharmacological unconsciousness (Greicius *et al.*, 2008; Boveroux *et al.*, 2010; Stamatakis *et al.*, 2010; Amico *et al.*, 2014). Similarly, the frontoparietal network, which has been linked to perceptual and somesthetic processing (Smith *et al.*, 2009; Laird *et al.*, 2011) and is considered critical for conscious reportable perception (Dehaene *et al.*, 2003), showed reductions in functional connectivity during sleep (Larson-Prior *et al.*, 2009; Samann *et al.*, 2011; Boly *et al.*, 2012) and anaesthesia (Boveroux *et al.*, 2010). The salience network, which has been involved in conflict monitoring, information integration, response selection, interoceptive processes (Seeley *et al.*, 2007; Smith *et al.*, 2009; Ploner *et al.*, 2010;

Wiech *et al.*, 2010) and the emotional counterpart of pain (Seeley *et al.*, 2007; Shackman *et al.*, 2011), also showed modulations in connectivity under propofol anaesthesia (Guldenmund *et al.*, 2013). Here, the positive correlation between CRS-R scores and the salience network anterior cingulate cortex could account for the preserved capacities of some patients to orient their attentional resources towards environmental salient stimuli, such as noxious stimulation, corroborating previous PET data (Boly *et al.*, 2008). With regards to sensory networks, little changes have been reported under physiological and pharmacological unconsciousness (Heine *et al.*, 2012). Nevertheless, propofol-induced disconnections have been shown between the default mode network and motor cortex, reticular activating system and the thalamus

(Stamatakis *et al.*, 2010). In particular, the thalamus is of critical importance to consciousness (Dehaene and Changeux, 2005; Tononi, 2008). In our analysis the significance of the thalamus was controlled by involving it among the regions of interest in the three large-scale networks, namely the default mode network, frontoparietal and salience. The direct comparison between patients in MCS and VS/UWS did not identify any differences in network-level thalamic connectivity. However, a recent study with patients with disorders of consciousness using a target-detection task showed that respondents had a greater connectivity between the anterior thalamus and prefrontal cortex. These findings suggest that thalamo-frontal circuits are important for cognitive top–down processing (Monti *et al.*, 2015). Interestingly, when the cerebellum was used as a control network, CRS-R total scores did not correlate with any regions of this network in patients. Such findings confirm previous suggestions that the cerebellum has minimal implication in conscious-related processing (Tononi, 2008; Yu *et al.*, 2015). Taken together, the positive correlation between clinical scores and each network's functional connectivity highlight that the here studied networks are an appropriate means to study, at least to a certain degree, residual cognitive function in this patient cohort.

Importantly for clinical practice, we further aimed at determining the capacity of each network to differentiate between patients in MCS and VS/UWS. In terms of functional MRI-based differentiation of patients, to date differences in functional connectivity have been observed only at the group-level for the default mode (Boly *et al.*, 2009; Vanhaudenhuyse *et al.*, 2010; Norton *et al.*, 2012; Soddu *et al.*, 2012; Demertzi *et al.*, 2014), the frontoparietal and the auditory networks (Demertzi *et al.*, 2014). Here, we replicated these findings and further showed group differences in functional connectivity for the salience, sensorimotor and visual networks. Moving towards single-patient network-based differentiation, we found that all networks were able to differentiate patients with an acceptable accuracy (>86%). Such high rate of accuracy can be partly attributed to the fact that the network ranking was based on features extracted from the same population for which between-group differences were already known. To avoid a double-dipping effect, we aimed at validating the most highly ranked network in two independently assessed patient data sets (Salzburg and New York) and across healthy controls. To that end, we opted for single-patient classification based on the connectivity strength of the auditory network. Based on this network's connectivity, 20 of the 22 new patients were classified congruently, i.e. the clinical diagnosis matched the classification outcome. Of note is that the classifier positioned the independently assessed patients closer to the decision plane compared to patients included in the training set. This could be explained by the abovementioned favouring of the Liège training data set during the network ranking procedure, which might have led to a stricter classification of the validation set. Although the intrinsic connectivity networks

have been shown to be robust independent of different scanning parameters (Van Dijk *et al.*, 2010), the different parameters employed in each of the three centres might also have influenced the classifier's estimation. Alternatively, the use of a relevance vector machine classifier (Phillips *et al.*, 2011), which returns probabilities of a patient belonging to a clinical condition instead of using a binary decision, could be a more sensitive way to classify patients less strictly.

The classification results further highlight the challenges posed by behavioural examination (Majerus *et al.*, 2005) which in many cases underestimates patients' level of consciousness (Schnakers *et al.*, 2009). Here, the validation of the auditory network's classifier worked congruently for the majority of the included patients (20/22). Interestingly, the patient who was misclassified as MCS had a profile of VS/UWS on the day of scan but evolved to MCS 38 days later. The other patient was misclassified as VS/UWS but had a clinical profile of MCS on the day of scanning based on the presence of localization to noxious stimulation (note that this behaviour could not be elicited in any other evaluations). The validation of the classifier's outcome to the clinical evaluation was used as a starting point in our analysis. Therefore, a well-defined diagnostic baseline was critical for the subsequent patient classification. To that end, repeated clinical examinations with the CRS-R (average number of assessments $n = 6$ per patient) were performed. The clinical diagnosis was further confirmed with FDG-PET imaging, which has been shown to have high sensitivity in identifying patients in MCS (Stender *et al.*, 2014). Therefore, patients with an ambiguous profile on clinical assessment and neuroimaging data were not included in the analysis. Similarly, patients who received sedatives to minimize motion in the scanner (Soddu *et al.*, 2011) were further excluded. The reason to exclude sedated patients was because of our limited understanding of the potential effect of anaesthetics on network connectivity (Heine *et al.*, 2012). We here recognize the importance of increasing the classification power for patients scanned after receiving anaesthetics, given that many patients undergo anaesthesia not only to restrict scanner motion but also for neuroprotective reasons (Schifilliti *et al.*, 2010). Future investigations which will aim to disentangle between the variances of anaesthetics and pathology in functional connectivity measures are certainly essential. Finally, even though patients were scanned in an 'awake' state, the monitoring of patients' state of vigilance during data acquisition was not feasible because of technical difficulties. Hence, one cannot exclude the possibility that patients could have fallen asleep during scanning, which could subsequently influence the assessment of functional connectivity.

One explanation of why the auditory network was identified as the system with the highest discriminative capacity could concern its underlying functional neuroanatomy. Apart from temporal cortices, the auditory network further encompasses regions in occipital cortex, pre- and

postcentral areas, insula and anterior cingulate cortex (Damoiseaux *et al.*, 2006; Smith *et al.*, 2009; Laird *et al.*, 2011; Maudoux *et al.*, 2012; Demertzi *et al.*, 2014). The direct comparison between patients in MCS and VS/UWS restricted the identified areas to bilateral auditory and visual cortices. This pattern of auditory-visual functional connectivity has been previously described in normal conscious subjects during rest as well (Eckert *et al.*, 2008) and is in line with functional MRI results in consciousness research. For example, preserved functional MRI activity in temporal and occipital areas has been shown for healthy subjects during mental counting of auditory temporal irregularities; interestingly, this activation was identified only in those subjects who were attentive and aware of the auditory violations (Bekinschtein *et al.*, 2009). At a functional level, the auditory-visual functional connectivity, also referred to as crossmodal interaction, is considered relevant for multisensory integration (Clavagnier *et al.*, 2004). Multisensory integration has been suggested as a facilitator for top–down influences of higher-order regions to create predictions of forthcoming sensory events (Engel *et al.*, 2001). Such top–down connectivity was recently found with an EEG oddball paradigm that differentiated patients in MCS from VS/UWS (Boly *et al.*, 2011). Interestingly, decreased crossmodal auditory-visual interaction has been reported in healthy subjects with preserved structural connections but under pharmacologically-induced anaesthesia (Boveroux *et al.*, 2010). In that study, recovery of consciousness paralleled the restoration of the crossmodal connectivity suggesting a critical role of this connectivity pattern to consciousness level-dependent states.

In our results, the crossmodal interaction was more preserved in patients in MCS compared to unresponsive patients. The reduction in functional connectivity between the auditory-visual cortices in VS/UWS could be partly attributed to disrupted anatomical connections, often encountered in post-comatose patients (Perlbarg *et al.*, 2009; Fernandez-Espejo *et al.*, 2010, 2011; Stevens *et al.*, 2014; van der Eerden *et al.*, 2014). The tight link between functional and structural connectivity was recently shown in primates during propofol-induced unconsciousness with regards to resting state functional MRI dynamic fluctuations. In this study, functional connectivity was fluctuating less frequently among distinct consciousness states, it was mostly linked to the state characterizing unconsciousness and this pattern was mostly explained by the underlying structural connectivity (Barttfeld *et al.*, 2015). Here, the negative differences between the two patient groups on voxel-based morphometry of grey and whiter matter segments is suggestive that the changes in functional connectivity cannot be fully attributed to the underlying anatomical abnormalities. We recognize that analyses with diffusion-weighted imaging and its relation to functional data would allow for more confident statements about residual functional connectivity in our clinical sample.

In conclusion, we here identified that systems-level resting state functional MRI showed consciousness-dependent breakdown not only for the default mode network but also for the frontoparietal, salience, auditory, sensorimotor and visual networks. Functional connectivity between auditory and visual cortices was the most sensitive feature to accurately discriminate single patients into the categories of MCS and VS/UWS. Our findings point to the significance of multisensory integration and top–down processes in consciousness seemingly supported by crossmodal connectivity. In the future, efforts need to be made to promote the feasibility of such a complex approach in the clinical setting and promote the clinical utility of the resting paradigm for single-patient diagnostics.

Acknowledgements

Funding

Supplementary material

Supplementary material is available at *Brain* online.

References

Achard S, Delon-Martin C, Vertes PE, Renard F, Schenck M, Schneider F, et al. Hubs of brain functional networks are radically reorganized in comatose patients. Proc Natl Acad Sci USA 2012; 109: 20608–13.

Achard S, Kremer S, Schenck M, Renard F, Ong-Nicolas C, Namer JI, et al. Global functional disconnections in post-anoxic coma patient. Neuroradiol J 2011; 24: 311–5.

Amico E, Gomez F, Di Perri C, Vanhaudenhuyse A, Lesenfants D, Boveroux P, et al. Posterior cingulate cortex-related co-activation patterns: a resting state FMRI study in propofol-induced loss of consciousness. PLoS One 2014; 9: e100012.

Bardin JC, Schiff ND, Voss HU. Pattern classification of volitional functional magnetic resonance imaging responses in patients with severe brain injury. Arch Neurol 2012; 69: 176–81.

Barttfeld P, Uhrig L, Sitt JD, Sigman M, Jarraya B, Dehaene S. Signature of consciousness in the dynamics of resting-state brain activity. Proc Natl Acad Sci USA 2015; 112: 887–92.

Behzadi Y, Restom K, Liau J, Liu TT. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 2007; 37: 90–101.

Bekinschtein TA, Dehaene S, Rohaut B, Tadel F, Cohen L, Naccache L. Neural signature of the conscious processing of auditory regularities. Proc Natl Acad Sci USA 2009; 106: 1672–7.

Bishop CM. (2006). Pattern Recognition and Machine Learning. Singapore: Springer.

Boly M, Faymonville M-E, Schnakers C, Peigneux P, Lambermont B, Phillips C, et al. Perception of pain in the minimally conscious state with PET activation: an observational study. Lancet Neurol 2008; 7: 1013–20.

Boly M, Garrido MI, Gosseries O, Bruno MA, Boveroux P, Schnakers C, et al. Preserved feedforward but impaired top-down processes in the vegetative state. Science 2011; 332: 858–62.

Boly M, Perlbarg V, Marrelec G, Schabus M, Laureys S, Doyon J, et al. Hierarchical clustering of brain activity during human nonrapid eye movement sleep. Proc Natl Acad Sci USA 2012; 109: 5856–61.

Boly M, Tshibanda L, Vanhaudenhuyse A, Noirhomme Q, Schnakers C, Ledoux D, et al. Functional connectivity in the default network during resting state is preserved in a vegetative but not in a brain dead patient. Hum Brain Mapp 2009; 30: 2393–400.

Boveroux P, Vanhaudenhuyse A, Bruno MA, Noirhomme Q, Lauwick S, Luxen A, et al. Breakdown of within- and between-network resting state functional magnetic resonance imaging connectivity during propofol-induced loss of consciousness. Anesthesiology 2010; 113: 1038–53.

Bruno MA, Soddu A, Demertzi A, Laureys S, Gosseries O, Schnakers C, et al. Disorders of consciousness: moving from passive to resting state and active paradigms. Cog Neurosci 2010; 1: 193–203.

Buckner RL, Andrews-Hanna JR, Schacter DL. The brain's default network: anatomy, function, and relevance to disease. Ann N Y Acad Sci 2008; 1124: 1–38.

Burges C. A tutorial on support vector machines for pattern recognition. Data Min Knowl Discov 1998; 2: 121–67.

Casali AG, Gosseries O, Rosanova M, Boly M, Sarasso S, Casali KR, et al. A theoretically based index of consciousness independent of sensory processing and behavior. Sci Transl Med 2013; 5: 198ra05.

Chai XJ, Castanon AN, Ongur D, Whitfield-Gabrieli S. Anticorrelations in resting state networks without global signal regression. Neuroimage 2012; 59: 1420–8.

Clavagnier S, Falchier A, Kennedy H. Long-distance feedback projections to area V1: implications for multisensory integration, spatial awareness, and visual consciousness. Cogn Affect Behav Neurosci 2004; 4: 117–26.

Cole DM, Smith SM, Beckmann CF. Advances and pitfalls in the analysis and interpretation of resting-state FMRI data. Front Syst Neurosci 2010; 4: 8.

Coleman MR, Rodd JM, Davis MH, Johnsrude IS, Menon DK, Pickard JD, et al. Do vegetative patients retain aspects of language comprehension? Evidence from fMRI. Brain 2007; 130: 2494–507.

Crone JS, Soddu A, Holler Y, Vanhaudenhuyse A, Schurz M, Bergmann J, et al. Altered network properties of the fronto-parietal network and the thalamus in impaired consciousness. Neuroimage Clin 2014; 4: 240–48.

Damoiseaux JS, Rombouts SA, Barkhof F, Scheltens P, Stam CJ, Smith SM, et al. Consistent resting-state networks across healthy subjects. Proc Natl Acad Sci USA 2006; 103: 13848–53.

De Martino F, Gentile F, Esposito F, Balsi M, Di Salle F, Goebel R, et al. Classification of fMRI independent components using IC-fingerprints and support vector machine classifiers. Neuroimage 2007; 34: 177–94.

Dehaene S, Changeux JP. Ongoing spontaneous activity controls access to consciousness: a neuronal model for inattentional blindness. PLoS Biol 2005; 3: e141.

Dehaene S, Sergent C, Changeux JP. A neuronal network model linking subjective reports and objective physiological data during conscious perception. Proc Natl Acad Sci USA 2003; 100: 8520–5.

Demertzi A, Gómez F, Crone JS, Vanhaudenhuyse A, Tshibanda L, Noirhomme Q, et al. Multiple fMRI system-level baseline connectivity is disrupted in patients with consciousness alterations. Cortex 2014; 52: 35–46.

Di HB, Yu SM, Weng XC, Laureys S, Yu D, Li JQ, et al. Cerebral response to patient's own name in the vegetative and minimally conscious states. Neurology 2007; 68: 895–9.

Eckert MA, Kamdar NV, Chang CE, Beckmann CF, Greicius MD, Menon V. A cross-modal system linking primary auditory and visual cortices: evidence from intrinsic fMRI connectivity analysis. Hum Brain Mapp 2008; 29: 848–57.

Edlow BL, Giacino JT, Wu O. Functional MRI and outcome in traumatic coma. Curr Neurol Neurosci Rep 2013; 13: 1–11.

Engel AK, Fries P, Singer W. Dynamic predictions: oscillations and synchrony in top-down processing. Nat Rev Neurosci 2001; 2: 704–16.

Fernandez-Espejo D, Bekinschtein T, Monti MM, Pickard JD, Junque C, Coleman MR, et al. Diffusion weighted imaging distinguishes the vegetative state from the minimally conscious state. Neuroimage 2011; 54: 103–12.

Fernandez-Espejo D, Junque C, Cruse D, Bernabeu M, Roig-Rovira T, Fabregas N, et al. Combination of diffusion tensor and functional magnetic resonance imaging during recovery from the vegetative state. BMC Neurol 2010; 10: 77.

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci USA 2005; 102: 9673–8.

Giacino JT, Ashwal S, Childs N, Cranford R, Jennett B, Katz DI, et al. The minimally conscious state: Definition and diagnostic criteria. Neurology 2002; 58: 349–53.

Giacino JT, Fins JJ, Laureys S, Schiff ND. Disorders of consciousness after acquired brain injury: the state of the science. Nat Rev Neurol 2014; 10: 99–114.

Giacino JT, Kalmar K, Whyte J. The JFK Coma Recovery Scale-Revised: measurement characteristics and diagnostic utility. Arch Phys Med Rehabi 2004; 85: 2020–9.

Greicius MD, Kiviniemi V, Tervonen O, Vainionpaa V, Alahuhta S, Reiss AL, et al. Persistent default-mode network connectivity during light sedation. Hum Brain Mapp 2008; 29: 839–47.

Greicius MD, Krasnow B, Reiss AL, Menon V. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc Natl Acad Sci USA 2003; 100: 253–8.

Guldenmund P, Demertzi A, Boveroux P, Boly M, Vanhaudenhuyse A, Bruno MA, et al. Thalamus, brainstem and salience network connectivity changes during propofol-induced sedation and unconsciousness. Brain Connect 2013; 3: 273–85.

Heine L, Soddu A, Gomez F, Vanhaudenhuyse A, Tshibanda L, Thonnard M, et al. Resting state networks and consciousness. Alterations of multiple resting state network connectivity in physiological, pharmacological and pathological consciousness states. Front Psychol 2012; 3: 1–12.

Horovitz SG, Braun AR, Carr WS, Picchioni D, Balkin TJ, Fukunaga M, et al. Decoupling of the brain's default mode network during deep sleep. Proc Natl Acad Sci USA 2009; 106: 11376–81.

Huang Z, Dai R, Wu X, Yang Z, Liu D, Hu J, et al. The self and its resting state in consciousness: an investigation of the vegetative state. Hum Brain Mapp 2014; 35: 1997–2008.

Jennett B, Plum F. Persistent vegetative state after brain damage. A syndrome in search of a name. Lancet 1972; 1: 734–7.

Kelly RE Jr, Alexopoulos GS, Wang Z, Gunning FM, Murphy CF, Morimoto SS, et al. Visual inspection of independent components:

defining a procedure for artifact removal from fMRI data. J Neurosci Methods 2010; 189: 233–45.

Laird AR, Fox PM, Eickhoff SB, Turner JA, Ray KL, McKay DR, et al. Behavioral interpretations of intrinsic connectivity networks. J Cogn Neurosci 2011; 23: 4022–37.

Larson-Prior LJ, Zempel JM, Nolan TS, Prior FW, Snyder AZ, Raichle ME. Cortical network functional connectivity in the descent to sleep. Proc Natl Acad Sci USA 2009; 106: 4489–94.

Laureys S, Celesia G, Cohadon F, Lavrijsen J, Leon-Carrrion J, Sannita WG, et al. Unresponsive wakefulness syndrome: a new name for the vegetative state or apallic syndrome. BMC Med 2010; 8: 68.

Majerus S, Gill-Thwaites H, Andrews K, Laureys S. Behavioral evaluation of consciousness in severe brain damage. Prog Brain Res 2005; 150: 397–413.

Maki-Marttunen V, Diez I, Cortes JM, Chialvo DR, Villarreal M. Disruption of transfer entropy and inter-hemispheric brain functional connectivity in patients with disorder of consciousness. Front Neuroinform 2013; 7: 24.

Maudoux A, Lefebvre P, Cabay JE, Demertzi A, Vanhaudenhuyse A, Laureys S, et al. Auditory resting-state network connectivity in tinnitus: a functional MRI study. PLoS One 2012; 7: e36222.

Monti MM, Rosenberg M, Finoia P, Kamau E, Pickard JD, Owen AM. Thalamo-frontal connectivity mediates top-down cognitive functions in disorders of consciousness. Neurology 2015; 84: 167–73.

Monti MM, Vanhaudenhuyse A, Coleman MR, Boly M, Pickard JD, Tshibanda L, et al. Willful modulation of brain activity in disorders of consciousness. N Engl J Med 2010; 362: 579–89.

Murphy K, Birn RM, Handwerker DA, Jones TB, Bandettini PA. The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? Neuroimage 2009; 44: 893–905.

Norton L, Hutchison RM, Young GB, Lee DH, Sharpe MD, Mirsattari SM. Disruptions of functional connectivity in the default mode network of comatose patients. Neurology 2012; 78: 175–81.

Ovadia-Caro S, Nir Y, Soddu A, Ramot M, Hesselmann G, Vanhaudenhuyse A, et al. Reduction in inter-hemispheric connectivity in disorders of consciousness. PLoS One 2012; 7: e37238.

Owen AM, Coleman MR, Boly M, Davis MH, Laureys S, Pickard JD. Detecting awareness in the vegetative state. Science 2006; 313: 1402.

Perlbarg V, Puybasset L, Tollard E, Lehericy S, Benali H, Galanaud D. Relation between brain lesion location and clinical outcome in patients with severe traumatic brain injury: a diffusion tensor imaging study using voxel-based approaches. Hum Brain Mapp 2009; 30: 3924–33.

Phillips CL, Bruno MA, Maquet P, Boly M, Noirhomme Q, Schnakers C, et al. "Relevance vector machine" consciousness classifier applied to cerebral metabolism of vegetative and locked-in patients. Neuroimage 2011; 56: 797–808.

Ploner M, Lee MC, Wiech K, Bingel U, Tracey I. Prestimulus functional connectivity determines pain perception in humans. Proc Natl Acad Sci USA 2010; 107: 355–60.

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. A default mode of brain function. Proc Natl Acad Sci USA 2001; 98: 676–82.

Rosanova M, Gosseries O, Casarotto S, Boly M, Casali AG, Bruno MA, et al. Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients. Brain 2012; 135: 1308–20.

Saad ZS, Gotts SJ, Murphy K, Chen G, Jo HJ, Martin A, et al. Trouble at rest: how correlation patterns and group differences become distorted after global signal regression. Brain Connect 2012; 2: 25–32.

Saeys Y, Inza I, Larranaga P. A review of feature selection techniques in bioinformatics. Bioinformatics 2007; 23: 2507–17.

Samann PG, Wehrle R, Hoehn D, Spoormaker VI, Peters H, Tully C, et al. Development of the brain's default mode network from wakefulness to slow wave sleep. Cereb Cortex 2011; 21: 2082–93.

Sanders RD, Tononi G, Laureys S, Sleigh JW. Unresponsiveness not equal unconsciousness. Anesthesiology 2012; 116: 946–59.

Schiff ND, Rodriguez-Moreno D, Kamal A, Kim KH, Giacino JT, Plum F, et al. fMRI reveals large-scale network activation in minimally conscious patients. Neurology 2005; 64: 514–23.

Schifilliti D, Grasso G, Conti A, Fodale V. Anaesthetic-related neuroprotection: intravenous or inhalational agents? CNS Drugs 2010; 24: 893–907.

Schnakers C, Vanhaudenhuyse A, Giacino JT, Ventura M, Boly M, Majerus S, et al. Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment. BMC Neurol 2009; 9: 35.

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, et al. Dissociable intrinsic connectivity networks for salience processing and executive control. J Neurosci 2007; 27: 2349–56.

Shackman AJ, Salomons TV, Slagter HA, Fox AS, Winter JJ, Davidson RJ. The integration of negative affect, pain and cognitive control in the cingulate cortex. Nat Rev Neurosci 2011; 12: 154–67.

Sitt JD, King JR, El Karoui I, Rohaut B, Faugeras F, Gramfort A, et al. Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state. Brain 2014; 137: 2258–70.

Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, et al. Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci USA 2009; 106: 13040–5.

Soddu A, Vanhaudenhuyse A, Bahri MA, Bruno MA, Boly M, Demertzi A, et al. Identifying the default-mode component in spatial IC analyses of patients with disorders of consciousness. Hum Brain Mapp 2012; 33: 778–96.

Soddu A, Vanhaudenhuyse A, Demertzi A, Bruno MA, Tshibanda L, Di H, et al. Resting state activity in patients with disorders of consciousness. Funct Neurol 2011; 26: 37–43.

Stamatakis EA, Adapa RM, Absalom AR, Menon DK. Changes in resting neural connectivity during propofol sedation. PLoS One 2010; 5: e14224.

Stender J, Gosseries O, Bruno MA, Charland-Verville V, Vanhaudenhuyse A, Demertzi A, et al. Diagnostic precision of PET imaging and functional MRI in disorders of consciousness: a clinical validation study. Lancet 2014; 384: 514–22.

Stevens RD, Hannawi Y, Puybasset L. MRI for coma emergence and recovery. Curr Opin Crit Care 2014; 20: 168–73.

Tononi G. Consciousness as integrated information: a provisional manifesto. Biol Bull 2008; 215: 216–42.

Tsai YH, Yuan R, Huang YC, Yeh MY, Lin CP, Biswal BB. Disruption of brain connectivity in acute stroke patients with early impairment in consciousness. Front Psychol 2014; 4: 956.

van der Eerden AW, Khalilzadeh O, Perlbarg V, Dinkel J, Sanchez P, Vos PE, et al. White matter changes in comatose survivors of anoxic ischemic encephalopathy and traumatic brain injury: comparative diffusion-tensor imaging study. Radiology 2014; 270: 506–16.

Van Dijk KR, Hedden T, Venkataraman A, Evans KC, Lazar SW, Buckner RL. Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. J Neurophysiol 2010; 103: 297–321.

Van Dijk KR, Sabuncu MR, Buckner RL. The influence of head motion on intrinsic functional connectivity MRI. Neuroimage 2012; 59: 431–38.

Vanhaudenhuyse A, Demertzi A, Schabus M, Noirhomme Q, Bredart S, Boly M, et al. Two distinct neuronal networks mediate the awareness of environment and of self. J Cogn Neurosci 2011; 23: 570–8.

Vanhaudenhuyse A, Noirhomme Q, Tshibanda LJ, Bruno MA, Boveroux P, Schnakers C, et al. Default network connectivity reflects the level of consciousness in non-communicative brain-damaged patients. Brain 2010; 133: 161–71.

Whitfield-Gabrieli S, Nieto-Castanon A. Conn: a functional connect-ivity toolbox for correlated and anticorrelated brain networks. Brain Connect 2012; 2: 125–41.

Wiech K, Lin CS, Brodersen KH, Bingel U, Ploner M, Tracey I. Anterior insula integrates information about salience into perceptual decisions about pain. J Neurosci 2010; 30: 16324–31.

Wong CW, Olafsson V, Tal O, Liu TT. Anti-correlated networks, global signal regression, and the effects of caffeine in resting-state functional MRI. Neuroimage 2012; 63: 356–64.

Yu F, Jiang QJ, Sun XY, Zhang RW. A new case of complete primary cerebellar agenesis: clinical and imaging findings in a living patient. Brain 2015; 138: e353.

References

[1] C. Di Perri, J. Annen, G. Antonopoulos, E. Amico, C. Cavaliere, and S. Laureys. *Measuring consciousness through imaging*. Springer, 2016.

[2] Quentin Noirhomme, Ralph Brecheisen, Damien Lesenfants, Georgios Antonopoulos, and Steven Laureys. "Look at my classifier's result": Disentangling unresponsive from (minimally) conscious patients. *NeuroImage*, 145:288–303, 2017.

[3] Athena Demertzi, Georgios Antonopoulos, Lizette Heine, Henning U. Voss, Julia Sophia Crone, Carlo de Los Angeles, Mohamed Ali Bahri, Carol Di Perri, Audrey Vanhaudenhuyse, Vanessa Charland-Verville, Martin Kronbichler, Eugen Trinka, Christophe Phillips, Francisco Gomez, Luaba Tshibanda, Andrea Soddu, Nicholas D. Schiff, Susan Whitfield-Gabrieli, and Steven Laureys. Intrinsic functional connectivity differentiates minimally conscious from unresponsive patients. *Brain*, 138(9):2619–2631, Sep 2015.

[4] F Plum and J B Posner. The diagnosis of stupor and coma. *Contemporary neurology series*, 10:1–286, 1972.

[5] Steven Laureys. The neural correlate of (un)awareness: lessons from the vegetative state. *Trends in Cognitive Sciences*, 9(12):556–559, Dec 2005.

[6] Audrey Vanhaudenhuyse, Athena Demertzi, Manuel Schabus, Quentin Noirhomme, Serge Bredart, Melanie Boly, Christophe Phillips, Andrea Soddu, Andre Luxen, Gustave Moonen, and Steven Laureys. Two distinct neuronal networks mediate the awareness of environment and of self. *Journal of Cognitive Neuroscience*, 23(3):570–578, Mar 2011.

[7] Steven Laureys, Gastone G Celesia, Francois Cohadon, Jan Lavrijsen, José León-Carrión, Walter G Sannita, Leon Sazbon, Erich Schmutzhard, Klaus R von Wild, Adam Zeman, Giuliano Dolce, and European Task Force on Disorders of Consciousness. Unresponsive wakefulness syndrome: a new name for the vegetative state or apallic syndrome. *BMC Medicine*, 8(1):68, Dec 2010.

[8] C Schnakers, C Chatelle, A Demertzi, S Majerus, and S Laureys. What about pain in disorders of consciousness? *AAPS J*, 14(3):437–444, 2012.

[9] A Zeman. Persistent vegetative state. *Lancet (London, England)*, 350(9080):795–9, Sep 1997.

[10] B Jennett and F Plum. Persistent vegetative state after brain damage. A syndrome in search of a name. *Lancet (London, England)*, 1(7753):734–7, Apr 1972.

[11] Joseph T Giacino, S Ashwal, N Childs, R Cranford, B Jennett, D I Katz, J P Kelly, J H Rosenberg, J Whyte, R D Zafonte, and N D Zasler. The minimally conscious state: definition and diagnostic criteria. *Neurology*, 58(3):349–53, Feb 2002.

[12] M A Bruno, F Pellas, J L Bernheim, D Ledoux, S Goldman, A Demertzi, S Majerus, A Vanhaudenhuyse, V Blandin, M Boly, P Boveroux, G Moonen, S Laureys, and C Schnakers. [Life with Locked-In syndrome]. *Revue medicale de Liege*, 63(5-6):445–51, 2008.

[13] Mélanie Boly, Marie-Elisabeth Faymonville, Caroline Schnakers, Philippe Peigneux, Bernard Lambermont, Christophe Phillips, Patrizio Lancellotti, Andre Luxen, Maurice Lamy, Gustave Moonen, Pierre Maquet, and Steven Laureys. Perception of pain in the minimally conscious state with PET activation: an observational study. *The Lancet Neurology*, 7(11):1013–1020, Nov 2008.

[14] Camille Chatelle, Aurore Thibaut, John Whyte, Marie Danièle De Val, Steven Laureys, and Caroline Schnakers. Pain issues in disorders of consciousness. *Brain Injury*, 28(9):1202–1208, Aug 2014.

[15] Frédéric Faugeras, Benjamin Rohaut, Mélanie Valente, Jacobo Sitt, Sophie Demeret, Francis Bolgert, Nicolas Weiss, Alexandra Grinea, Clémence Marois, Marion Quirins, Athena Demertzi, Federico Raimondo, Damien Galanaud, Marie-Odile Habert, Denis Engemann, Louis Puybasset, and Lionel Naccache. Survival and consciousness recovery are better in the minimally conscious state than in the vegetative state. *Brain Injury*, 32(1):72–77, Jan 2018.

[16] Joseph T. Giacino, Kathleen Kalmar, and John Whyte. The JFK Coma Recovery Scale-Revised: Measurement characteristics and diagnostic utility11No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit upon the authors or upon any organization with which the authors are associated. *Archives of Physical Medicine and Rehabilitation*, 85(12):2020–2029, Dec 2004.

[17] Eelco F M Wijdicks, William R Bamlet, Boby V Maramattom, Edward M Manno, and Robyn L McClelland. Validation of a new coma scale: The FOUR score. *Annals of neurology*, 58(4):585–93, Oct 2005.

[18] G Teasdale, G Murray, L Parker, and B Jennett. Adding up the Glasgow Coma Score. *Acta neurochirurgica. Supplementum*, 28(1):13–6, 1979.

[19] Caroline Schnakers, Steve Majerus, Joseph Giacino, Audrey Vanhaudenhuyse, Marie-Aurelie Bruno, Melanie Boly, Gustave Moonen, Pierre Damas, Bernard Lambermont, Maurice Lamy, FranÇois Damas, Manfredi Ventura, and Steven Laureys. A French validation study of the Coma Recovery Scale-Revised (CRS-R). *Brain Injury*, 22(10):786–792, Jan 2008.

[20] Caroline Schnakers, Audrey Vanhaudenhuyse, Joseph Giacino, Manfredi Ventura, Melanie Boly, Steve Majerus, Gustave Moonen, and Steven Laureys. Diagnostic accuracy of the vegetative and minimally conscious state: Clinical consensus versus standardized neurobehavioral assessment. *BMC Neurology*, 9(1):35, Dec 2009.

[21] Olivia Gosseries, Marie-Aurélie Bruno, Camille Chatelle, Audrey Vanhaudenhuyse, Caroline Schnakers, Andrea Soddu, and Steven Laureys. Disorders of consciousness: what's in a name? *NeuroRehabilitation*, 28(1):3–14, 2011.

[22] Damian Cruse, Ithabi Gantner, Andrea Soddu, and Adrian M. Owen. Lies, damned lies and diagnoses: Estimating the clinical utility of assessments of covert awareness in the vegetative state. *Brain Injury*, 28(9):1197–1201, Aug 2014.

[23] P Guldenmund, J Stender, L Heine, and S Laureys. Mindsight: diagnostics in disorders of consciousness. *Critical care research and practice*, 2012:624724, Nov 2012.

[24] F. Bloch. Nuclear Induction. *Physical Review*, 70(7-8):460–474, Oct 1946.

[25] S Ogawa, T M Lee, A R Kay, and D W Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci U S A*, 87(24):9868–9872, Dec 1990.

[26] A K Shukla and Utham Kumar. Positron emission tomography: An overview. *Journal of medical physics*, 31(1):13–21, Jan 2006.

[27] Daniel C. Sullivan, Nancy A. Obuchowski, Larry G. Kessler, David L. Raunig, Constantine Gatsonis, Erich P. Huang, Marina Kondratovich, Lisa M. McShane, Anthony P. Reeves, Daniel P. Barboriak, Alexander R. Guimaraes, Richard L. Wahl, and For the RSNA-QIBA Metrology Working Group. Metrology Standards for Quantitative Imaging Biomarkers. *Radiology*, 277(3):813, 2015.

[28] John Ashburner and Karl J. Friston. Unified segmentation. *NeuroImage*, 26(3):839–851, Jul 2005.

[29] B Biswal, Yetkin, Haughton, and Hyde. - Functional connectivity in the motor cortex of resting human brain using. *Magn Reson Med*, 34(9):537–541, 1995.

[30] M. Boly, M.R. Coleman, M.H. Davis, A. Hampshire, D. Bor, G. Moonen, P.A. Maquet, J.D. Pickard, S. Laureys, and A.M. Owen. When thoughts become action: An fMRI paradigm to study volitional brain activity in non-communicative brain injured patients. *NeuroImage*, 36(3):979–992, Jul 2007.

[31] Stanislas Dehaene and Jean-Pierre Changeux. Ongoing Spontaneous Activity Controls Access to Consciousness: A Neuronal Model for Inattentional Blindness. *PLoS Biology*, 3(5):e141, Apr 2005.

[32] M E Raichle, A M MacLeod, A Z Snyder, W J Powers, D A Gusnard, and G L Shulman. A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2):676–82, Jan 2001.

[33] M. D. Greicius, B. Krasnow, A. L. Reiss, and V. Menon. Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences*, 100(1):253–258, Jan 2003.

[34] Lucina Q. Uddin and Lucina Q. Uddin. Anatomy of the Salience Network. *Salience Network of the Human Brain*, pages 5–10, Jan 2017.

[35] William W Seeley, Vinod Menon, Alan F Schatzberg, Jennifer Keller, Gary H Glover, Heather Kenna, Allan L Reiss, and Michael D Greicius. Dissociable intrinsic connectivity networks for salience processing and executive control. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 27(9):2349–56, Feb 2007.

[36] Michael W. Cole and Walter Schneider. The cognitive control network: Integrated cortical regions with dissociable functions. *NeuroImage*, 37(1):343–360, Aug 2007.

[37] M. Ploner, M. C. Lee, K. Wiech, U. Bingel, and I. Tracey. Prestimulus functional connectivity determines pain perception in humans. *Proceedings of the National Academy of Sciences*, 107(1):355–360, Jan 2010.

[38] K. Wiech, C.-s. Lin, K. H. Brodersen, U. Bingel, M. Ploner, and I. Tracey. Anterior Insula Integrates Information about Salience into Perceptual Decisions about Pain. *Journal of Neuroscience*, 30(48):16324–16331, Dec 2010.

[39] Miranda Scolari, Katharina N Seidl-Rathkopf, and Sabine Kastner. Functions of the human frontoparietal attention network: Evidence from neuroimaging. *Current Opinion in Behavioral Sciences*, 1:32–39, Feb 2015.

[40] Scott Marek and Nico U F Dosenbach. The frontoparietal network: function, electrophysiology, and importance of individual precision mapping. *Dialogues in clinical neuroscience*, 20(2):133–140, 2018.

[41] Angela R. Laird, P. Mickle Fox, Simon B. Eickhoff, Jessica A. Turner, Kimberly L. Ray, D. Reese McKay, David C. Glahn, Christian F. Beckmann, Stephen M. Smith, and Peter T. Fox. Behavioral Interpretations of Intrinsic Connectivity Networks. *Journal of Cognitive Neuroscience*, 23(12):4022–4037, Dec 2011.

[42] Lizette Heine, Andrea Soddu, Francisco Gómez, Audrey Vanhaudenhuyse, Luaba Tshibanda, Marie Thonnard, Vanessa Charland-Verville, Murielle Kirsch, Steven Laureys, and Athena Demertzi. Resting State Networks and Consciousness. *Frontiers in Psychology*, 3:295, 2012.

[43] Lizette Heine, Maïté Castro, Charlotte Martial, Barbara Tillmann, Steven Laureys, and Fabien Perrin. Exploration of Functional Connectivity During Preferred Music Stimulation in Patients with Disorders of Consciousness. *Frontiers in Psychology*, 6:1704, Nov 2015.

[44] N Nakayama, A Okumura, J Shinoda, T Nakashima, and T Iwama. Relationship between regional cerebral metabolism and consciousness disturbance in traumatic diffuse brain injury without large focal lesions: an FDG-PET study with statistical parametric mapping analysis. *Journal of neurology, neurosurgery, and psychiatry*, 77(7):856–62, Jul 2006.

[45] I. Astola, D. Escudero, L. Forcelledo, L. Viña, C. Vigil, and F. González. The role of 18 F-fluorodeoxyglucose PET/CT in ruling out vegetative state. *Medicina Intensiva (English Edition)*, 41(2):127–129, Mar 2017.

[46] David E. Levy, John J. Sidtis, David A. Rottenberg, Jens O. Jarden, Stephen C. Strother, Vijay Dhawan, James Z. Ginos, Mark J. Tramo, Alan C. Evans, and Fred Plum. Differences in cerebral blood flow and glucose utilization in vegetative versus locked-in patients. *Annals of Neurology*, 22(6):673–682, Dec 1987.

[47] J Rudolf, M Ghaemi, M Ghaemi, W F Haupt, B Szelies, and W D Heiss. Cerebral glucose metabolism in acute and persistent vegetative state. *Journal of neurosurgical anesthesiology*, 11(1):17–24, Jan 1999.

[48] C Tommasino, C Grana, G Lucignani, G Torri, and F Fazio. Regional cerebral metabolism of glucose in comatose and vegetative state patients. *Journal of neurosurgical anesthesiology*, 7(2):109–16, Apr 1995.

[49] S Laureys, C Lemaire, P Maquet, C Phillips, and G Franck. Cerebral metabolism during vegetative state and after recovery to consciousness. *Journal of neurology, neurosurgery, and psychiatry*, 67(1):121, Jul 1999.

[50] Johan Stender, Kristian Nygaard Mortensen, Aurore Thibaut, Sune Darkner, Steven Laureys, Albert Gjedde, and Ron Kupers. The Minimal Energetic Requirement of Sustained Awareness after Brain Injury. *Current Biology*, 26(11):1494–1499, Jun 2016.

[51] Johan Stender, Olivia Gosseries, Marie-Aurélie Bruno, Vanessa Charland-Verville, Audrey Vanhaudenhuyse, Athena Demertzi, Camille Chatelle, Marie Thonnard, Aurore Thibaut, Lizette Heine, Andrea Soddu, Mélanie Boly, Caroline Schnakers, Albert Gjedde, and Steven Laureys. Diagnostic precision of PET imaging and functional MRI in disorders of consciousness: a clinical validation study. *The Lancet*, 384(9942):514–522, Aug 2014.

[52] A. M. Owen, M. R. Coleman, M. Boly, M. H. Davis, S. Laureys, and J. D. Pickard. Detecting Awareness in the Vegetative State. *Science*, 313(5792):1402–1402, Sep 2006.

[53] D Rodriguez Moreno, N D Schiff, J Giacino, K Kalmar, and J Hirsch. A network approach to assessing cognition in disorders of consciousness. *Neurology*, 75(21):1871–8, Nov 2010.

[54] Tristan Andres Bekinschtein, Facundo Francisco Manes, Mirta Villarreal, Adrian Mark Owen, and Valeria Della-Maggiore. Functional imaging reveals movement preparatory activity in the vegetative state. *Frontiers in human neuroscience*, 5:5, 2011.

[55] Martin M. Monti, Audrey Vanhaudenhuyse, Martin R. Coleman, Melanie Boly, John D. Pickard, Luaba Tshibanda, Adrian M. Owen, and Steven Laureys. Willful Modulation of Brain Activity in Disorders of Consciousness. *New England Journal of Medicine*, 362(7):579–589, Feb 2010.

[56] H. B. Di, S. M. Yu, X. C. Weng, S. Laureys, D. Yu, J. Q. Li, P. M. Qin, Y. H. Zhu, S. Z. Zhang, and Y. Z. Chen. Cerebral response to patient's own name in the vegetative and minimally conscious states. *Neurology*, 68(12):895–899, Mar 2007.

[57] S. Laureys, S. Goldman, C. Phillips, P. Van Bogaert, J. Aerts, A. Luxen, G. Franck, and P. Maquet. Impaired Effective Cortical Connectivity in Vegetative State: Preliminary Investigation Using PET. *NeuroImage*, 9(4):377–382, Apr 1999.

[58] Steven Laureys and Nicholas D. Schiff. Coma and consciousness: Paradigms (re)framed by neuroimaging. *NeuroImage*, 61(2):478–491, Jun 2012.

[59] B Beuthien-Baumann, W Handrick, T Schmidt, W Burchert, L Oehme, J Kropp, G Schackert, J Pinkert, and W-G Franke. Persistent vegetative state: evaluation of brain metabolism and brain perfusion with PET and SPECT. *Nuclear medicine communications*, 24(6):643–9, Jun 2003.

[60] S. Silva, X. Alacoque, O. Fourcade, K. Samii, P. Marque, R. Woods, J. Mazziotta, F. Chollet, and I. Loubinoux. Wakefulness and loss of awareness: Brain and brainstem interaction in the vegetative state. *Neurology*, 74(4):313–320, Jan 2010.

[61] Marie Aurélie Bruno, Audrey Vanhaudenhuyse, Aurore Thibaut, Gustave Moonen, and Steven Laureys. From unresponsive wakefulness to minimally conscious PLUS and functional locked-in syndromes: Recent advances in our understanding of disorders of consciousness. *Journal of Neurology*, 258(7):1373–1384, 2011.

[62] Cristina Rosazza, Adrian Andronache, Davide Sattin, Maria Grazia Bruzzone, Giorgio Marotta, Anna Nigri, Stefania Ferraro, Davide Rossi Sebastiano, Luca Porcu, Anna Bersano, Riccardo Benti, Matilde Leonardi, Ludovico D'Incerti, and Ludovico Minati. Multimodal study of default-mode network integrity in disorders of consciousness. *Annals of Neurology*, 79(5):841–853, May 2016.

[63] Smadar Ovadia-Caro, Yuval Nir, Andrea Soddu, Michal Ramot, Guido Hesselmann, Audrey Vanhaudenhuyse, Ilan Dinstein, Jean-Flory L Tshibanda, Melanie Boly, Michal Harel, Steven Laureys, and Rafael Malach. Reduction in interhemispheric connectivity in disorders of consciousness. *PloS one*, 7(5):e37238, 2012.

[64] Armin Iraji, Randall R Benson, Robert D Welch, Brian J O'Neil, John L Woodard, Syed Imran Ayaz, Andrew Kulek, Valerie Mika, Patrick Medado, Hamid Soltanian-Zadeh, Tianming Liu, E Mark Haacke, and Zhifeng Kou. Resting State Functional Connectivity in Mild Traumatic Brain Injury at the Acute Stage: Independent Component and Seed-Based Analyses. *Journal of neurotrauma*, 32(14):1031–45, Jul 2015.

[65] S Laureys, ME Faymonville, A Luxen, M Lamy, G Franck, and P Maquet. Restoration of thalamocortical connectivity after recovery from persistent vegetative state. *The Lancet*, 355(9217):1790–1791, May 2000.

[66] Steven Laureys, Sylvie Antoine, Melanie Boly, Sandra Elincx, Marie-Elisabeth Faymonville, Jacques Berré, Bernard Sadzot, Martine Ferring, Xavier De Tiège, Patrick van Bogaert, Isabelle Hansen, Pierre Damas, Nicolas Mavroudakis, Bernard Lambermont, Guy Del Fiore, Joël Aerts, Christian Degueldre, Christophe Phillips, George Franck, Jean-Louis Vincent, Maurice Lamy, André Luxen, Gustave Moonen, Serge Goldman, and Pierre Maquet. Brain function in the vegetative state. *Acta neurologica Belgica*, 102(4):177–85, Dec 2002.

[67] Bernard J Baars, Thomas Z Ramsøy, and Steven Laureys. Brain, conscious experience and the observing self. *Trends in neurosciences*, 26(12):671–5, Dec 2003.

[68] Stanislas Dehaene and Lionel Naccache. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, 79(1-2):1–37, Apr 2001.

[69] Geraint Rees, Gabriel Kreiman, and Christof Koch. Neural correlates of consciousness in humans. *Nature Reviews Neuroscience*, 3(4):261–270, 2002.

[70] Pierre Boveroux, Vincent Bonhomme, Mélanie Boly, Audrey Vanhaudenhuyse, Pierre Maquet, and Steven Laureys. Brain Function in Physiologically, Pharmacologically, and Pathologically Altered States of Consciousness. *International Anesthesiology Clinics*, 46(3):131–146, 2008.

[71] Pierre Maquet. Understanding non rapid eye movement sleep through neuroimaging. *The World Journal of Biological Psychiatry*, 11(sup1):9–15, Jan 2010.

[72] Leo Breiman. Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). *Statistical Science*, 16(3):199–231, Aug 2001.

[73] Danilo Bzdok. Classical Statistics and Statistical Learning in Imaging Neuroscience. *Frontiers in Neuroscience*, 11:543, Oct 2017.

[74] Danilo Bzdok, Naomi Altman, and Martin Krzywinski. Statistics versus machine learning. *Nature Methods*, 15(4):233–234, Apr 2018.

[75] Galit Shmueli. To Explain or to Predict? *Statistical Science*, 25(3):289–310, Aug 2010.

[76] K. J. Friston, C. D. Frith, P. F. Liddle, and R. S. J. Frackowiak. Functional Connectivity: The Principal-Component Analysis of Large (PET) Data Sets. *Journal of Cerebral Blood Flow & Metabolism*, 13(1):5–14, Jan 1993.

[77] Karl J. Friston. Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*, 2(1-2):56–78, Jan 1994.

[78] Mark Jenkinson, Christian F. Beckmann, Timothy E.J. Behrens, Mark W. Woolrich, and Stephen M. Smith. FSL. *NeuroImage*, 62(2):782–790, Aug 2012.

[79] B. J. Winer, Donald R. Brown, and Kenneth M. Michels. *Statistical principles in experimental design*. McGraw-Hill, New York :, 3rd ed. edition, 1991.

[80] K. J. (Karl J.) Friston, John Ashburner, Stefan Kiebel, Thomas Nichols, and William D. Penny. *Statistical parametric mapping : the analysis of funtional brain images*. Elsevier/Academic Press, 2007.

[81] Jija S James, Pg Rajesh, Anuvitha Vs Chandran, and Chandrasekharan Kesavadas. fMRI paradigm designing and post-processing tools. *The Indian journal of radiology & imaging*, 24(1):13–21, Jan 2014.

[82] B M de Jong, A T Willemsen, and A M Paans. Regional cerebral blood flow changes related to affective speech presentation in persistent vegetative state. *Clinical neurology and neurosurgery*, 99(3):213–6, Aug 1997.

[83] Fatima Siddiqui and Qazi Mazhar Ali. Performance of non-parametric classifiers on highly skewed data, 2016.

[84] Yugal Kumar and G. Sahoo. Analysis of Parametric & Non Parametric Classifiers for Classification Technique using WEKA. *International Journal of Information Technology and Computer Science*, 4(7):43–49, Jul 2012.

[85] Bernhard E. Boser, Bernhard E. Boser, Isabelle M. Guyon, and Vladimir N. Vapnik. A Training Algorithm for Optimal Margin Classifiers. *PROCEEDINGS OF THE 5TH ANNUAL ACM WORKSHOP ON COMPUTATIONAL LEARNING THEORY*, pages 144—-152, 1992.

[86] Vladimir Naumovich. Vapnik. *The nature of statistical learning theory*. Springer, 2000.

[87] Sergios Theodoridis and Konstantinos Koutroumbas. *Pattern recognition*. Academic Press, 2009.

[88] Christopher M. Bishop. *Pattern recognition and machine learning*. Springer, 2006.

[89] Bo-Yu Chu, Chia-Hua Ho, Cheng-Hao Tsai, Chieh-Yen Lin, and Chih-Jen Lin. Warm Start for Parameter Selection of Linear Classifiers. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '15*, pages 149–158, New York, New York, USA, 2015. ACM Press.

[90] Janaina Mourao-Miranda, Emanuelle Reynaud, Francis McGlone, Gemma Calvert, and Michael Brammer. The impact of temporal compression and space selection on SVM analysis of single-subject and multi-subject fMRI data. *Neuroimage*, 33(4):1055–1065, Dec 2006.

[91] Demis Hassabis, Carlton Chu, Geraint Rees, Nikolaus Weiskopf, Peter D Molyneux, and Eleanor A Maguire. Decoding neuronal ensembles in the human hippocampus. *Current biology : CB*, 19(7):546–54, Apr 2009.

[92] Jessica Schrouff, Caroline Kussé, Louis Wehenkel, Pierre Maquet, and Christophe Phillips. Decoding Semi-Constrained Brain Activity from fMRI Using Support Vector Machines and Gaussian Processes. *PLoS ONE*, 7(4):e35860, Apr 2012.

[93] Asa Ben-Hur and Jason Weston. A User's Guide to Support Vector Machines. In *Data Mining Techniques for the Life Sciences*, pages 223–239. Humana Press, 2010.

[94] Thomas Hofmann, Bernhard Schölkopf, and Alexander J. Smola. Kernel methods in machine learning. *The Annals of Statistics*, 36(3):1171–1220, Jun 2008.

[95] Carl Edward. Rasmussen and Christopher K. I. Williams. *Gaussian processes for machine learning*. MIT Press, 2006.

[96] Pierre Geurts, Damien Ernst, and Louis Wehenkel. Extremely randomized trees. *Machine Learning*, 63(1):3–42, Apr 2006.

[97] S. B. Kotsiantis. Decision trees: a recent overview. *Artificial Intelligence Review*, 39(4):261–283, Apr 2013.

[98] Laura Elena Raileanu and Kilian Stoffel. Theoretical comparison between the Gini Index and Information Gain criteria. *Annals of Mathematics and Artificial Intelligence*, 41(1):77–93, May 2004.

[99] Gilles Louppe. *Understanding Random Forests: From Theory to Practice*. PhD thesis, University of Liege, Jul 2014.

[100] Tin Kam Ho. The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8):832–844, 1998.

[101] Leo Breiman. Random Forests. *Machine Learning*, 45(1):5–32, 2001.

[102] Leo Breiman. Bagging Predictors. *Machine Learning*, 24:123–140, 1996.

[103] B. Efron. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1):1–26, Jan 1979.

[104] Nikolaus Kriegeskorte, W Kyle Simmons, Patrick S F Bellgowan, and Chris I Baker. Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience*, 12(5):535–540, May 2009.

[105] Bradley Efron and Robert Tibshirani. Improvements on Cross-Validation: The 632+ Bootstrap Method. *Journal of the American Statistical Association*, 92(438):548–560, Jun 1997.

[106] Joset A. Etzel, Valeria Gazzola, and Christian Keysers. An introduction to anatomical ROI-based fMRI classification analysis. *Brain Research*, 1282:114–125, Jul 2009.

[107] Steven Lemm, Benjamin Blankertz, Thorsten Dickhaus, and Klaus-Robert Müller. Introduction to machine learning for brain imaging. *NeuroImage*, 56(2):387–399, May 2011.

[108] Choong-Wan Woo, Luke J Chang, Martin A Lindquist, and Tor D Wager. Building better biomarkers: brain models in translational neuroimaging. *Nature neuroscience*, 20(3):365–377, Feb 2017.

[109] Gavin C. Cawley and Nicola L. C. Talbot. On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation. *Journal of Machine Learning Research*, 11(Jul):2079–2107, 2010.

[110] Rajeev Kumar and Abhaya Indrayan. Receiver operating characteristic (ROC) curve for medical researchers. *Indian Pediatrics*, 48(4):277–287, Apr 2011.

[111] Kevin Markham. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874, Dec 2006.

[112] J. Kent Martin, J. Kent Martin, and D. S. Hirschberg. Small Sample Statistics for Classification Error Rates II: Confidence Intervals and Significance Tests. *DEPT. OF INF. AND COMP. SCI*, 1996.

[113] D. Berrar, I. Bradbury, and W. Dubitzky. Avoiding model selection bias in small-sample genomic datasets. *Bioinformatics*, 22(10):1245–1250, May 2006.

[114] Quentin Noirhomme, Damien Lesenfants, Francisco Gomez, Andrea Soddu, Jessica Schrouff, Gaëtan Garraux, André Luxen, Christophe Phillips, and Steven Laureys. Biased binomial assessment of cross-validated estimation of classification accuracies illustrated in diagnosis predictions. *NeuroImage: Clinical*, 4:687–694, 2014.

[115] Denis A Engemann, Federico Raimondo, Jean-Rémi King, Benjamin Rohaut, Gilles Louppe, Frédéric Faugeras, Jitka Annen, Helena Cassol, Olivia Gosseries, Diego Fernandez-Slezak, Steven Laureys, Lionel Naccache, Stanislas Dehaene, and Jacobo D Sitt. Robust EEG-based cross-site and cross-protocol classification of states of consciousness. *Brain*, 141(11):3179–3192, Nov 2018.

[116] Christophe L. Phillips, Marie-Aurelie Bruno, Pierre Maquet, Mélanie Boly, Quentin Noirhomme, Caroline Schnakers, Audrey Vanhaudenhuyse, Maxime Bonjean, Roland Hustinx, Gustave Moonen, André Luxen, and Steven Laureys. "Relevance vector machine" consciousness classifier applied to cerebral metabolism of vegetative and locked-in patients. *NeuroImage*, 56(2):797–808, May 2011.

[117] Athena Demertzi, Francisco Gómez, Julia Sophia Crone, Audrey Vanhaudenhuyse, Luaba Tshibanda, Quentin Noirhomme, Marie Thonnard, Vanessa Charland-Verville, Murielle Kirsch, Steven Laureys, and Andrea Soddu. Multiple fMRI system-level baseline connectivity is disrupted in patients with consciousness alterations. *Cortex*, 52:35–46, Mar 2014.

[118] Xuehai Wu, Qihong Zou, Jin Hu, Weijun Tang, Ying Mao, Liang Gao, Jianhong Zhu, Yi Jin, Xin Wu, Lu Lu, Yaojun Zhang, Yao Zhang, Zhengjia Dai, Jia-Hong Gao, Xuchu Weng, Liangfu Zhou, Georg Northoff, Joseph T Giacino, Yong He, and Yihong Yang. Intrinsic Functional Connectivity Patterns Predict Consciousness Level and Recovery Outcome in Acquired Brain Injury. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 35(37):12932–46, Sep 2015.

[119] Ming Song, Yi Yang, Jianghong He, Zhengyi Yang, Shan Yu, Qiuyou Xie, Xiaoyu Xia, Yuanyuan Dang, Qiang Zhang, Xinhuai Wu, Yue Cui, Bing Hou, Ronghao Yu, Ruxiang Xu, and Tianzi Jiang. Prognostication of chronic disorders of consciousness using brain functional networks and clinical characteristics. *eLife*, 7, 2018.

[120] M. Dash and H. Liu. Feature selection for classification. *Intelligent Data Analysis*, 1(1-4):131–156, Jan 1997.

[121] Jianyu Miao and Lingfeng Niu. A Survey on Feature Selection. *Procedia Computer Science*, 91:919–926, Jan 2016.

[122] Isabelle Guyon and André Elisseeff. An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research*, 3(Mar):1157–1182, 2003.

[123] Isabelle Guyon and André Elisseeff. *Journal of machine learning research : JMLR.*, volume 3. MIT Press, 2001.

[124] Esra Mahsereci Karabulut, Selma Ayşe Özel, and Turgay İbrikçi. A comparative study on the effect of feature selection on classification accuracy. *Procedia Technology*, 1:323–327, Jan 2012.

[125] Athena Demertzi, Jacobo Diego Sitt, Simone Sarasso, and Wim Pinxten. Measuring states of pathological (un)consciousness: research dimensions, clinical applications, and ethics†. *Neuroscience of Consciousness*, 2017(1), Jan 2017.

[126] Mario Rosanova, Olivia Gosseries, Silvia Casarotto, Mélanie Boly, Adenauer G. Casali, Marie-Aurélie Bruno, Maurizio Mariotti, Pierre Boveroux, Giulio Tononi, Steven Laureys, and Marcello Massimini. Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients. *Brain*, 135(4):1308–1320, Apr 2012.

[127] Adenauer G. Casali, Olivia Gosseries, Mario Rosanova, Mélanie Boly, Simone Sarasso, Karina R. Casali, Silvia Casarotto, Marie Aurélie Bruno, Steven Laureys, Giulio Tononi, and Marcello Massimini. A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine*, 5(198), 2013.

[128] Jacobo Diego Sitt, Jean-Remi King, Imen El Karoui, Benjamin Rohaut, Frederic Faugeras, Alexandre Gramfort, Laurent Cohen, Mariano Sigman, Stanislas Dehaene, and Lionel Naccache. Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state. *Brain*, 137(8):2258–2270, Aug 2014.

[129] Denis A Engemann, Federico Raimondo, Jean-Rémi King, Benjamin Rohaut, Gilles Louppe, Frédéric Faugeras, Jitka Annen, Helena Cassol, Olivia Gosseries, Diego Fernandez-Slezak, Steven Laureys, Lionel Naccache, Stanislas Dehaene, and Jacobo D Sitt. Robust EEG-based cross-site and cross-protocol classification of states of consciousness. *Brain*, Oct 2018.

[130] Jonathan C. Bardin, Nicholas D. Schiff, and Henning U. Voss. Pattern Classification of Volitional Functional Magnetic Resonance Imaging Responses in Patients With Severe Brain Injury. *Archives of Neurology*, 69(2):176, Feb 2012.

[131] T Bekinschtein, R Leiguarda, J Armony, A Owen, S Carpintiero, J Niklison, L Olmos, L Sigman, and F Manes. Emotion processing in the minimally conscious state. *Journal of neurology, neurosurgery, and psychiatry*, 75(5):788, May 2004.

[132] M. C. Rousseau, S. Confort-Gouny, A. Catala, J. Graperon, J. Blaya, E. Soulier, P. Viout, D. Galanaud, Y. Le Fur, P. J. Cozzone, and J. P. Ranjeva. A MRS-MRI-fMRI exploration of the brain. Impact of long-lasting persistent vegetative state. *Brain Injury*, 22(2):123–134, Jan 2008.

[133] M A Ramsay, T M Savege, B R Simpson, and R Goodwin. Controlled sedation with alphaxalone-alphadolone. *British medical journal*, 2(5920):656–9, Jun 1974.

[134] John Ashburner. A fast diffeomorphic image registration algorithm. *NeuroImage*, 38(1):95–113, Oct 2007.

[135] Koene R.A. Van Dijk, Mert R. Sabuncu, and Randy L. Buckner. The influence of head motion on intrinsic functional connectivity MRI. *NeuroImage*, 59(1):431–438, Jan 2012.

[136] Yashar Behzadi, Khaled Restom, Joy Liau, and Thomas T Liu. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, 37(1):90–101, Aug 2007.

[137] Xiaoqian J Chai, Alfonso Nieto Castañón, Dost Ongür, and Susan Whitfield-Gabrieli. Anticorrelations in resting state networks without global signal regression. *NeuroImage*, 59(2):1420–8, Jan 2012.

[138] Michael D Fox, Abraham Z Snyder, Justin L Vincent, Maurizio Corbetta, David C Van Essen, and Marcus E Raichle. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27):9673–8, Jul 2005.

[139] Susan Whitfield-Gabrieli and Alfonso Nieto-Castanon. <i>Conn</i> : A Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks. *Brain Connectivity*, 2(3):125–141, Jun 2012.

[140] Pierre Boveroux, Audrey Vanhaudenhuyse, Marie-Aurélie Bruno, Quentin Noirhomme, Séverine Lauwick, André Luxen, Christian Degueldre, Alain Plenevaux, Caroline Schnakers, Christophe Phillips, Jean-François Brichant, Vincent Bonhomme, Pierre Maquet, Michael D. Greicius, Steven Laureys, and Mélanie Boly. Breakdown of within- and between-network Resting State Functional Magnetic Resonance Imaging Connectivity during Propofol-induced Loss of Consciousness. *Anesthesiology*, 113(5):1038–1053, Nov 2010.

[141] M. De Luca, C.F. Beckmann, N. De Stefano, P.M. Matthews, and S.M. Smith. fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *NeuroImage*, 29(4):1359–1367, Feb 2006.

[142] Damien A. Fair, Alexander L. Cohen, Jonathan D. Power, Nico U. F. Dosenbach, Jessica A. Church, Francis M. Miezin, Bradley L. Schlaggar, and Steven E. Petersen. Functional Brain Networks Develop from a "Local to Distributed" Organization. *PLoS Computational Biology*, 5(5):e1000381, May 2009.

[143] Audrey Maudoux, Philippe Lefebvre, Jean-Evrard Cabay, Athena Demertzi, Audrey Vanhaudenhuyse, Steven Laureys, and Andrea Soddu. Auditory Resting-State Network Connectivity in Tinnitus: A Functional MRI Study. *PLoS ONE*, 7(5):e36222, May 2012.

[144] Marcus E Raichle. The restless brain. *Brain connectivity*, 1(1):3–12, 2011.

[145] Moriah E Thomason, Emily L Dennis, Anand A Joshi, Shantanu H Joshi, Ivo D Dinov, Catie Chang, Melissa L Henry, Rebecca F Johnson, Paul M Thompson, Arthur W Toga, Gary H Glover, John D Van Horn, and Ian H Gotlib. Resting-state fMRI can reliably map neural networks in children. *NeuroImage*, 55(1):165–75, Mar 2011.

[146] Giulio Tononi. Consciousness as integrated information: a provisional manifesto. *The Biological bulletin*, 215(3):216–42, Dec 2008.

[147] Wilson Yu and Esther Krook-Magnuson. Cognitive Collaborations: Bidirectional Functional Connectivity Between the Cerebellum and the Hippocampus. *Frontiers in systems neuroscience*, 9:177, 2015.

[148] Y. Saeys, I. Inza, and P. Larranaga. A review of feature selection techniques in bioinformatics. *Bioinformatics*, 23(19):2507–2517, Oct 2007.

[149] Wei Zhu, Xuena Wang, Yeming Ma, Manlong Rao, James Glimm, and John S Kovach. Detection of cancer-specific markers amid massive mass spectral data. *Proceedings of the National Academy of Sciences of the United States of America*, 100(25):14666–71, Dec 2003.

[150] Ian C. Gould, Alana M. Shepherd, Kristin R. Laurens, Murray J. Cairns, Vaughan J. Carr, and Melissa J. Green. Multivariate neuroanatomical classification of cognitive subtypes in schizophrenia: A support vector machine learning approach. *NeuroImage: Clinical*, 6:229–236, Jan 2014.

[151] Chih-chung Chang, Chih-chung Chang, and Chih-Jen Lin. LIBSVM: a Library for Support Vector Machines. *Transactions in Intelligent Systems and Technology*, 2001.

[152] Pegah Kassraian-Fard, Caroline Matthis, Joshua H. Balsters, Marloes H. Maathuis, and Nicole Wenderoth. Promises, Pitfalls, and Basic Guidelines for Applying Machine Learning Classifiers to Psychiatric Imaging Data, with Autism as an Example. *Frontiers in Psychiatry*, 7:177, Dec 2016.

[153] Simon Clavagnier, Arnaud Falchier, and Henry Kennedy. Long-distance feedback projections to area V1: implications for multisensory integration, spatial awareness, and visual consciousness. *Cognitive, affective & behavioral neuroscience*, 4(2):117–26, Jun 2004.

[154] Andreas K. Engel, Pascal Fries, and Wolf Singer. Dynamic predictions: Oscillations and synchrony in top–down processing. *Nature Reviews Neuroscience*, 2(10):704–716, Oct 2001.

[155] T. A. Bekinschtein, S. Dehaene, B. Rohaut, F. Tadel, L. Cohen, and L. Naccache. Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences*, 106(5):1672–1677, 2009.

[156] Steven Laureys, Adrian M Owen, and Nicholas D Schiff. Brain function in coma, vegetative state, and related disorders. *The Lancet Neurology*, 3(9):537–546, Sep 2004.

[157] A Thibaut, M Bruno, C Chatelle, O Gosseries, A Vanhaudenhuyse, A Demertzi, C Schnakers, M Thonnard, V Charland-Verville, C Bernard, M Bahri, C Phillips, M Boly, R Hustinx, and S Laureys. Metabolic activity in external and internal awareness networks in severely brain-damaged patients. *Journal of Rehabilitation Medicine*, 44(6):487–494, 2012.

[158] Johan Stender, Ron Kupers, Anders Rodell, Aurore Thibaut, Camille Chatelle, Marie-Aurélie Bruno, Michael Gejl, Claire Bernard, Roland Hustinx, Steven Laureys, and Albert Gjedde. Quantitative Rates of Brain Glucose Metabolism Distinguish Minimally Conscious from Vegetative State Patients. *Journal of Cerebral Blood Flow & Metabolism*, 35(1):58–65, Jan 2015.

[159] Brent A. Vogt. Posterior cingulate, precuneal and retrosplenial cortices: cytology and components of the neural network correlates of consciousness. *Progress in Brain Research*, 150:205–217, Jan 2005.

[160] Athena Demertzi, Andrea Soddu, and Steven Laureys. Consciousness supporting networks. *Current Opinion in Neurobiology*, 23(2):239–244, Apr 2013.

[161] J Ashburner and K J Friston. Nonlinear spatial normalization using basis functions. *Human brain mapping*, 7(4):254–66, 1999.

[162] Jenny Crinion, John Ashburner, Alex Leff, Matthew Brett, Cathy Price, and Karl Friston. Spatial normalization of lesioned brains: performance evaluation and impact on fMRI analyses. *NeuroImage*, 37(3):866–75, Sep 2007.

[163] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot. Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage*, 15(1):273–289, Jan 2002.

[164] Edmund T. Rolls, Marc Joliot, and Nathalie Tzourio-Mazoyer. Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. *NeuroImage*, 122:1–5, Nov 2015.

[165] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(Oct):2825–2830, 2011.

[166] XULEI YANG, QING SONG, and YUE WANG. A weighted support vector machine for data classification. *International Journal of Pattern Recognition and Artificial Intelligence*, 21(05):961–976, Aug 2007.

[167] V. I. Gorodetskiy and S. V. Serebryakov. Methods and algorithms of collective recognition. *Automation and Remote Control*, 69(11):1821–1851, Nov 2008.

[168] Sanjay Yadav and Sanyam Shukla. Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification. In *2016 IEEE 6th International Conference on Advanced Computing (IACC)*, pages 78–83. IEEE, Feb 2016.

[169] Mingrui Xia, Jinhui Wang, and Yong He. BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics. *PLoS ONE*, 8(7):e68910, Jul 2013.

[170] Melanie Boly, Marcello Massimini, Naotsugu Tsuchiya, Bradley R Postle, Christof Koch, and Giulio Tononi. Are the Neural Correlates of Consciousness in the Front or in the Back of the Cerebral Cortex? Clinical and Neuroimaging Evidence. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 37(40):9603–9613, Oct 2017.

[171] Michał Woźniak and Manuel Graña. A survey of multiple classifier systems as hybrid systems. *Information Fusion*, 16:3–17, Mar 2014.

[172] Matthew Brett, Alexander P. Leff, Chris Rorden, and John Ashburner. Spatial Normalization of Brain Images with Focal Lesions Using Cost Function Masking. *NeuroImage*, 14(2):486–500, Aug 2001.

[173] Nikos K. Logothetis and Brian A. Wandell. Interpreting the BOLD Signal. *Annual Review of Physiology*, 66(1):735–769, Mar 2004.

[174] Nikos K. Logothetis. What we can do and what we cannot do with fMRI. *Nature*, 453(7197):869–878, Jun 2008.

[175] Arne Ekstrom. How and when the fMRI BOLD signal relates to underlying neural activity: the danger in dissociation. *Brain research reviews*, 62(2):233–44, Mar 2010.

[176] Chase R. Figley and Patrick W. Stroman. The role(s) of astrocytes and astrocyte activity in neurometabolism, neurovascular coupling, and the production of functional neuroimaging signals. *European Journal of Neuroscience*, 33(4):577–588, Feb 2011.

[177] Eduardo R Zimmer, Maxime J Parent, Débora G Souza, Antoine Leuzy, Clotilde Lecrux, Hyoung-Ihl Kim, Serge Gauthier, Luc Pellerin, Edith Hamel, and Pedro Rosa-Neto. [18F]FDG PET signal is driven by astroglial glutamate transport. *Nature Neuroscience*, 20(3):393–395, Mar 2017.

[178] Tram Nguyen, Christina Baun, and Poul Flemming Høilund-Carlsen. An account of data entry inconsistencies and their impact on positron emission tomography quantification. *Medicine*, 97(37):e12312, Sep 2018.

[179] Adeline Lo, Herman Chernoff, Tian Zheng, and Shaw-Hwa Lo. Why significant variables aren't automatically good predictors. *Proceedings of the National Academy of Sciences of the United States of America*, 112(45):13892–7, Nov 2015.

[180] Danilo Bzdok, Denis Engemann, Olivier Grisel, Gaël Varoquaux, and Bertrand Thirion. Prediction and inference diverge in biomedicine: Simulations and real-world data. *bioRxiv*, page 327437, May 2018.

[181] Fumitoshi Kodaka, Hiroshi Ito, Miho Shidahara, Harumasa Takano, Hidehiko Takahashi, Ryosuke Arakawa, Kazuhiko Nakayama, and Tetsuya Suhara. Positron emission tomography inter-scanner differences in dopamine D2 receptor binding measured with [11C]FLB457. *Annals of Nuclear Medicine*, 24(9):671–677, Nov 2010.

[182] Jessica Schrouff, J. M. Monteiro, L. Portugal, M. J. Rosa, C. Phillips, and J. Mourão-Miranda. Embedding Anatomical or Functional Knowledge in Whole-Brain Multiple Kernel Learning Models. *Neuroinformatics*, 16(1):117–143, Jan 2018.

[183] Jitka Annen, Gianluca Frasso, Julia Sophia Crone, Lizette Heine, Carol Di Perri, Charlotte Martial, Helena Cassol, Athena Demertzi, Lionel Naccache, and Steven Laureys. Regional brain volumetry and brain function in severely brain-injured patients. *Annals of Neurology*, 83(4):842–853, Apr 2018.

Acknowledgements

It is very difficult to acknowledge all those people who have provided an input to complete the PhD. Might have been a smile or a nice word in a moment of frustration. Sharing a coffee on a Saturday evening or during one of the holidays I spent in the lab. Those people took away a lot of pressure, which was making my head boiling sometimes, but is not easy to recall their names. Some of those names I do not even know. Besides them though, there are those people that had to deal with me systematically inside or outside the lab. No matter whether it was financial, scientific, psychological, material or even negative input, their support has been determinant for fulfilling the PhD. Unfortunately, for this part there are no references to run through, but I will try to remember all of them. My apologies though, if I forget someone. It is July 22, 3am and I just finalized my thesis - meaning that my brain is working in "safe mode".

I would like to start this acknowledgement part with my father, my grandfather and my grandmother, who were there in the beginning of this journey, but did not make it to see the end. Πατέρα, εντάξει ο "σκληρός ο δίσκος", μαλάκωσε. Παππούλη, δεν ξέρω αν τα έφτιαξα τα κεφάλια αλλά ένα λιθαράκι πιστεύω να το έβαλα. Κ εσύ γιαγιάκα μη φοβάσαι, μια δουλούλα θα τη βρω πια.

Many thanks to my supervisors. At first, Prof. Steven Laureys, who believed in me and gave me this opportunity. His support, both on a scientific and personal level during my PhD were crucial for completing this work. Next, I would like to thank Prof Christophe Phillips, who provided me with all necessary methodological knowledge and advice to overcome all the obstacles that showed up.

Special thanks to those people that were able to cope with me at work but also in their free time. First, Athena Demertzi who was the one who motivated me to come and supported me during the whole period of my studies, and my coding guru Seven (or Fede for most of the people). Holy Jitka that always appeared like a "deus ex machina" and Gianluca -the voice of god when it comes to statistics. Vinci (without comments cos everyone in Liege knows him), Erik (RTFM), Lizette (no more excuses, I'll finish it now), Dr. Prof. "Γιατί είσαι τόσο πολύ..." Enrico, my future neighbors Olivia and Zig. Thank you people for the great moments during work, for dinners, brainstorming and beerstorming, video games etc.