Aus dem Institut/der Klinik für Psychiatrie und Psychotherapie, Campus Mitte
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Diagnostic classification and relapse prediction in alcohol dependence using fMRI

From classification algorithm to imaging approach

zur Erlangung des akademischen Grades
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Quoc Phoi Dam
aus Vietnam

Datum der Promotion: 22.06.2014

TABLE OF CONTENTS

ZUSAMMENFASSUNG

Trotz zahlreicher Hinweise darauf, dass die zerebralen Aktivierungsmusterin der funktionellen Magnetresonanztomographie (fMRI) in Reaktion auf krankheitsassoziierte Stimuli zur Diagnostik und Prognose verwendet werden könnten, wird das fMRI zur Bestimmung von Biomarkern der Alkoholabhängigkeit in der Praxis bisher nicht angewendet. Das Ziel dieser Dissertation war die Entwicklung von Voraussetzungen, die die Identifizierung von Alkoholabhängigkeit und auch die Vorhersage des Rückfallrisikos in der klinischen Praxis mittels fMRI ermöglicht. Diese Arbeit beinhaltet (1) die Identifizierung wichtiger Hirnregionen (ROI; region of interest) im Prozess der diagnostischen und prognostischen Klassifikation von fMRI; (2) die Anwendung der Bildgebung und (3) die Validierung der Methode.

Die erste Analyse in dieser Dissertation fokussiert auf die Identifizierbarkeit von Hirnregionen (ROIs), die für die Klassifikation bedeutsam sind. Diese Studie wurde an 50 alkoholkranken Patienten und 57 gesunden Kontrollen durchgeführt. Die Ergebnisse zeigten die Überlegenheit der Güte der diagnostischen Klassifikation (Patienten vs Gesunde) mittels funktioneller ROIs z.B. für das ventrale Striatum (VS, 63.9% Genauigkeit), das vorderer Cingulum (ACC, 62.8% Genauigkeit) im Vergleich zur Klassifikationsgenauigkeit mittels der Gesamthirndaten (61.8% Genauigkeit) oder des präfrontalen Cortex (PFC, 51.8% Genauigkeit). Diese Daten legen die praktische Anwendbarkeit von funktionellen ROI Analysen auf das fMRI mit Hilfe multivariaten Methoden wie Support Vector MachineVerfahren (SVM) nahe.

Die zweite Analyse bezieht sich auf die Anwendbarkeit der Methode auf die Vorhersage eine Trinkrückfalls. Diese Studie wurde bei 40 Patienten, aufgeteilt in 20 abstinente und 20 rückfällige Patienten durchgeführt. Die Patienten wurden zufällig aus den 50 alkoholkranken Patienten in der ersten Studie ausgewählt und nach der Entgiftung über einen sechs monatigen Verlauf nachuntersucht. Die Klassifikationsergebnisse zeigten, dass die Aktivität des VS, des ACC und der Insula eine hohe Genauigkeit in der Rückfallvorhersage mit 63.7%, 58.1% und 71.5% besitzen. Hier beizeigten das rechte VS und das rechte ACC höhere prädiktive Werte als dieselben Strukturen in der linken Hemisphäre (75.9% und 68.2% im Vergleich zu 53.1% und 58.9%). Eine Kombination aus dem rechten VS, dem rechten ACC und der bilateralen Insula ergab eine bessere Vorhersage (76.9% Genauigkeit, $p<0.0001$).

Die dritte Analyse fokussiert auf die Anwendung der Bildgebungsverfahren und verwendet die Daten aus der zweiten Studie. Die Methode basiert auf einem Ranking-Index, dem Grad der Aktivierungsunterschiede zwischen den zu trennenden Klassen. Die Ergebnisse zeigten eine gute Reliabilität und Genauigkeit des Index welche durch hohe Konvergenz und deren hoher Korrelation mit den Ergebnissen der SVM Klassifikatoren charakterisiert ist. Weiterhin erreicht die Rückfallvorhersage für den Patienteneine Genauigkeit von 80%, 72.5% und 70% (p=0.00002, p=0.0011 und p=0.0032), wenn die Vorhersage auf den Ranking-Indizes der Aktivierungsmuster des rechten VS, rechten ACC oder der bilateralen Insula basiert.

Zur Überprüfung und Validierung des Klassifikationsansatzes auch in der klinischen Praxis wurden zwei Pilot-Analysen durchgeführt. Basis dieser Analysen waren die Daten der dritten Studie. Basis dieser Analysen waren die Daten der dritten Studie. Die erste Pilotanalyse umfasste das Monitoring des Krankheitsverlaufes nach Entzug mittels der spektralen Darstellung der zerebralen Aktivierungen. Es zeigte sich ein signifikanter Unterschied in den Spektren des VS beim Vergleich der Patienten mit und ohne Trinkrückfall. Die zweite Pilot-Analyse zielte auf das Erfassen on korrelativen Zusammenhängen zwischen Bildgebung und klinischen Parametern ab mit dem Ziel einer Validierung an den Verhaltensdaten der Patienten. Die Ergebnisse zeigten eine mittelgradige Korrelation zwischen dem Ranking-Index und dem durch eine visuelle Analogskala gemessenen Grad von Durst und Hunger (VAS-TH) auf der Basis Aktivierungsdaten des rechten VS, des rechten ACC und der bilateralen Insula (z. B. für die Insula, R=-0.674, p=0.003).

Trotz einiger methodischer Limitationen zeigen die vorgestellten Daten die Relevanz bestimmter Hirnregionen für die Diagnostik und die Vorhersage des Verlaufes bei Alkoholabhängigkeit mit Hilfe des fMRI. Die Daten sind eine erste Grundlage für die weitere Forschung zur Frage inwieweit fMRI basierte Biomarker bei der Diagnostik und Prognose neuropsychiatrischer Störungen eine klinische Bedeutung erlangen kann.

Keywords: *Alkoholabhängigkeit, Rückfallvorhersage, fMRI, SVM, ROI, ROI-Kombination, Bayes-Inferenz, Erkennbarkeit Ebene.*

ABSTRACT

Although there is much evidence indicating that cerebral activation patterns in response to disease-related stimuli measured by functional Magnetic Resonance Imaging (fMRI) may be used as criteria for diagnosis as well as prognosis, the application of fMRI as biomarkers in alcohol dependence remains challenging. The aim of this dissertation was to develop a framework which enables the identification of alcohol dependence as well as the prediction of relapse risk in clinical practice using fMRI, namely (1) Specifying important brain regions in fMRI classification; (2) Approaching imaging; (3) Validating the approach.

The first analysis in this dissertation focused on the identifiability of important brain regions for the classification. This study was conducted on 50 alcoholic patients and 57 healthy controls. The results showed the outperformance of diagnostic classification (patient vs. healthy) on the activation images of functional regions of interest (ROIs) collected from important brain structures in alcohol dependence, e.g. from the ventral striatum (VS, 63.9% accuracy); the anterior cingulate cortex (ACC, 62.8% accuracy) compared to those from the whole brain (61.8%, accuracy); the prefrontal cortex (PFC, 51.8% accuracy). The evidence suggests the practicality of functional ROI analyses in fMRI classification using multivariate methods such as support vector machine (SVM).

The second analysis referred to the applicability of such an approach to the relapse prediction. This study was conducted on 40 patients including 20 relapsers and 20 abstainers drawn randomly from the 50 alcoholic patients used in the first study and followed up six months after detoxification. The results showed that the prediction using the activation images of VS, ACC and insula achieved high accuracies (63.7%, 58.1% and 71.5%, respectively). In addition, the activation images of VS and ACC recorded in the right hemisphere were more predictive than those in the left hemisphere (75.9% and 68.2% vs. 53.1% and 58.9% accuracy, respectively); and a combination of the individual predictions from these ROIs including the right VS, right ACC and bilateral insula gave a better prediction (76.9% accuracy; $p<0.0001$).

The third analysis offered an imaging approach. This study was conducted using the data of the second study. The method was centered on the ranking index characterizing the degree of separation of activation images between the two classes investigated. The results showed

reliability and certainty of the index through the characteristics of convergence and the strong and positive correlation between it and outputs of the SVM classifiers. Further, based on the ranking indices of the activation images of the right VS, right ACC and bilateral insula, the relapse prediction for the patients achieved 80%, 72.5% and 70% accuracy, respectively (p=0.00002, p=0.0011 and p=0.0032).

In order to examine applicability of the approach in clinical practice, the two pilot analyses were conducted on the data of the third study. The first pilot analysis involved the monitoring of disease progression after withdrawal using spectral representation of the cerebral activations. The results showed a significant difference in the spectrum of activation images of the VS when comparing the patients with and without drinking relapse. The second pilot analysis was captured on correlative relationships between imaging and clinical variables with the aim of validating the data on the behaviour of patients, which can make an inference of the analyzed brain disorder more reliable. The results disclosed a moderate correlation between the ranking index and the visual analog rating scale of thirst and hunger (VAS-TH) on the basis of activation data of the right VS, the right ACC and bilateral insula (e.g. for the insula, R=-0.674; p=0.003).

Despite several methodological limitations, the presented data show the relevance of specific brain regions to the diagnosis and prediction of the progression of alcohol dependence using fMRI. The data are the first basis for further research on the question of whether fMRI-based biomarkers can attain a clinical significance in the diagnosis and prognosis of neuropsychiatric disorders.

Keywords: *Alcohol dependence, relapse prediction, fMRI, SVM, ROI, ROI combination, Bayesian inference, discernibility level*.

List of abbreviations

ACC	Anterior cingulate cortex
ADS	Alcohol dependent score
AUQ	Alcohol urge questionnaire
BOLD	Blood oxygen level dependent
CV	Cross validation
Dec	Decision value
DSM-IV	Diagnostic and statistical manual of mental disorders
DS	Dorsal striatum
GABA	Gamma-aminobutyric acid
ICD-10	International statistical classification of diseases and related health problems
fMRI	Functional magnetic resonance imaging
LTD	Long-term depression
LTP	Long-term potentiation
MNI	Montreal neurological institute
mPFC	Medial prefrontal cortex
NAc	Nucleus accumbens
NMDA	N-methyl-D-aspartate
OCDS	Obsessive compulsive drinking scale
OFC	Orbital frontal cortex
PFC	Prefrontal cortex
RI	Ranking index
ROI	Region of interest
sd	Standard deviation
sRI	Ranking index for subject
SNR	Signal-noise ratio
SVM	Support vector machine
VAS-TH	Visual analog rating scale of thirst and hunger
VS	Ventral striatum
VTA	Ventral tegmental area

List of figures

List of tables

CHAPTER I

INTRODUCTION

Since its discovery by ancient Egypt and Greece (5th Before Christ), alcohol has been seen as a "drink madness" substance, and drunkenness has been referred to as a body and soul sickness (William et al., 2001). Along the time line, together with the advancement of science and technology, many mysteries of alcohol addiction have been gradually uncovered. Nowadays, alcohol addiction or alcohol dependence, originated from long-term alcohol drinking, is recognized as a common neurobiological brain disorder, which is treatable (Helga, 2011). The source of its pathogenesis comes not only from alcohol but also from many factors such as genetics, environment, stress, personality, comorbidity, drug history, and so on. It eventually leads to neuroadaptation to the effects of alcohol (Koob & Le Moal, 2008). The structural change of the brain in adapting to environmental factors is a natural characteristic (Jones and Bonci, 2005), and the characteristics of brain activity at a given time can reflect the condition of alcohol-dependent patient at that time (De Witte, 2004; Koob & Volkow, 2010). However, at present the evaluation of such a condition is based mostly on clinical manifestations through direct physical examination. Although there are significant improvements in clinical consultation, the accuracy of diagnosis is much dependent on subjective measures of physicians and patients. Therefore, a more objective and accurate method is a practical need in the treatment and follow-up of alcohol-dependent patient. With the aid of functional magnetic resonance imaging (fMRI) and the methods of data analysis, this has gradually become achievable. A specific question posed here was whether fMRI can provide useful biomarkers in clinical practice for diagnosis as well as prediction of the relapse risk after detoxification, and this was also the problem that we aimed to address.

BACKGROUND

ALCOHOL DEPENDENCE

Alcohol abuse and alcohol dependence are significant public health problems all over the world. With the serious medical, economic and social consequences, the World Health Organization

(WHO) has viewed them as one of the leading risk factors for premature death and disabilities in the world, which is in the same order as tobacco and hypertension (Helga, 2011).

Alcohol is a toxic substance in all aspects of its direct and indirect effects on a wide range of body organs and systems (Rehm et al., 2009). The effects of alcohol cause medical, psychological and social damage. As the toxic effects of alcohol damage all organs of the body, excessive alcohol use has serious health consequences to the individual and may lead to liver cirrhosis, gastritis, ulcer, pancreatitis, gastrointestinal cancers, neuropsychiatric diseases, cardiovascular diseases, etc. (Room et al., 2005; Mack et al., 2010). With chronic drinking and repeated intoxication, a cluster of interrelated behavioural, physical and cognitive symptoms develops which is referred to as alcohol dependence (Thomas et al., 2001).

What is alcohol dependence?

Alcohol dependence, also known as alcohol addiction, is a chronically relapsing disorder characterized by criteria such as tolerance development, withdrawal symptoms, drug craving and reduced control of drug intake (WHO, 1992; Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV; (American Psychiatric Association (APA), 1994) and its Text Revision (DSM-IV-TR; APA, 2000); Table 1.1).

Table 1.1. **DSM-IV-TR diagnostic criteria for alcohol dependence**

A maladaptative pattern of alcohol use, leading to clinically significant impairment or distress, as manifested by three (or more) of the following, occurring at any time in the same 12-month period:

(1) Tolerance, as defined by either of the following:

(a) A need for markedly increased amounts of the alcohol to achieve intoxication or desired effect

(b) Markedly diminished effect with continued use of the same amount of the alcohol

(2) Withdrawal, as manifested by either of the following:

(a) The characteristic withdrawal syndrome for the alcohol

(b) Alcohol is taken to relieve or avoid withdrawal symptoms

(3) Alcohol is often taken in larger amounts or over a longer period than was intended

(4) There is persistent desire or unsuccessful efforts to cut down or control alcohol use

(5) A great deal of time is spent in activities necessary to obtain the alcohol (e.g. driving long distances), use alcohol or recover from its effects

(6) Important social, occupational, or recreational activities are given up or reduced because of alcohol use

(7) The alcohol use is continued despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by the substance (e.g. continued drinking despite that an ulcer was made worse by alcohol consumption).

- Criteria (1), (2) may describe the physical dependence.
- Criteria (3), (4) may describe the state of 'craving', which is a strong desire and urge to consume alcohol, as well as loss of control.
- Criteria (5), (6), (7) refer to the compulsive state and reflect the social and medical consequences of alcohol consumption.

Although the clinical criteria were established in DSM-IV or in several questionnaire protocols such as Alcohol Dependence Scale (ADS), Michigan Alcoholism Screening Test (MAST), Alcohol Urge Questionnaire (AUQ), Obsessive Compulsive Drinking Scale (OCDS), etc. with the aim of supporting the diagnosis of alcohol dependent condition more accurately, clinicians often don't have clear boundaries to diagnose definitely the condition of the disease (Mack et al., 2010; Helga, 2011). This suggests a need to develop better support tools in the future.

Stages of addiction

Drug addiction, including alcohol addiction, is today seen as a chronic relapsing condition characterized by (a) compulsion to seek and take the drug, (b) loss of control in limiting intake, and (c) emergence of a negative emotional state (e.g. dysphoria, anxiety, irritability) when access to the drug is prevented (Koob & Le Moal, 2005). The chronic effects of alcohol cause neuroadaptation in brain structure, plasticity and altered gene expression, leading to persistent changes in brain functions and transition from controlled to compulsive alcohol use (Helga, 2011). Such an addiction cycle is composed of three stages: 'binge/intoxication', 'withdrawal/negative affect', and 'preoccupation/anticipation' (craving) (Koob & Volkow, 2010).

The stage of 'binge/intoxication': *VTA and VS including nucleus accumbens*

This stage is characterized by a positively reinforcing effect, primarily mediated by the mesolimbic dopamine system, and is an important starting point for the transition to addiction (Koob & Volkow, 2010). The mesolimbic dopamine system plays a core role in reward, and the initial action of alcohol reward has been hypothesized to be dependent on dopamine release in this system (Heinz et al., 2009). Alcohol, via endorphin release in the ventral tegmental area (VTA), stimulates inhibitory opioid receptors located on GABAergic interneurons in the VTA and thereby indirectly disinhibits dopamine neurons (Fig. 1.1) (Steven et al., 2006). On the other

hand, the nucleus accumbens is located strategically (Fig. 1.4) to receive important information of the limbic system from the amygdala, frontal cortex, and hippocampus which can be converted to motivational action through its connections with the extrapyramidal motor system. Thus, the nucleus accumbens plays a critical role in the acute reinforcing effects of drugs, together with the supporting role for the central nucleus of the amygdala (CeA) and ventral pallidum (Fig. 1.4) (Koob & Volkow, 2010).

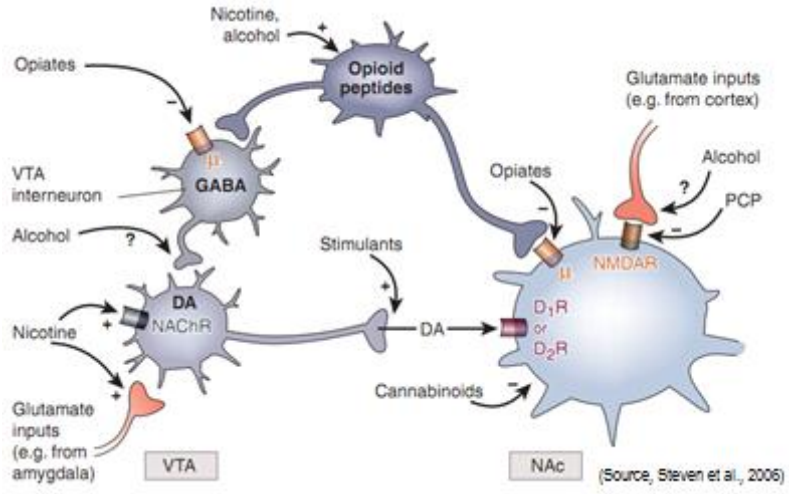


Figure 1.1. **Actions of opiates, nicotine, alcohol, and phencyclidine (PCP) in reward circuits.**

The dopamine neurons in ventral tegmental area (VTA) (bottom left) project to the nucleus accumbens (NAc) (bottom right). Different interneurons interact with VTA neurons and NAc neurons. Alcohol, acting on $GABA_A$ receptors in the VTA, can cause dopamine release (Source, Steven et al., 2006).

The stage of 'withdrawal/negative affect': *the extended amygdala*

The stage of acute withdrawal is characterized by changes of the within-system changes reflected by a decrease of dopaminergic activity in the mesolimbic dopamine system and by the between-system recruitment of neurotransmitter systems that convey stress and anxiety-like effects such as corticotropin-releasing factor (CRF) and dynorphin (Koob & Le Moal, 2008).

Within-system neuroadaptations

A within-system neuroadaptation in addiction is a molecular or cellular change within the reward circuit in order to adapt to overactivity of hedonic processing associated with addiction, which results in a decrease in reward function (Koob & Volkow, 2010). Decreases in activity of the mesolimbic dopamine system and decreases in serotonergic neurotransmission in the nucleus accumbens was recorded during alcohol withdrawal in a study on rats (Weiss et al, 1996): "Withdrawal from the chronic ethanol diet produces a progressive suppression in the release of dopamine and serotonergic neurotransmitters in the nucleus accumbens over the 8 hour

withdrawal period. Self-administration of ethanol reinstates and maintains brain dopamine release at pre-withdrawal levels." In addition, many studies of neurochemicals as well as imaging have shown that long-lasting reduction in the numbers of dopamine D2 receptors reflecting a hypodopaminergic state and the hypoactivity of the orbitofrontal-infralimbic cortex system in drug abusers compared with controls during this time (Volkow et al., 2003).

Between-system neuroadaptations: mutual changes between reward system and antireward system

In addiction, a between-system neuroadaptation is a circuitry change where the antireward circuit (brain stress circuit) is activated by excessive activity of the reward circuit. This activation generates opposing actions to limit the reward function (Koob & Le Moal, 2008). Both the hypothalamic–pituitary–adrenal axis (HPA) and the brain stress/aversive system mediated by the corticotropin-releasing factor (CRF) are activated during acute withdrawal from chronic administration of all addictive drugs with a common response of increasing adrenocorticotropic hormone, corticosterone and CRF (Koob & Kreek, 2007). Simultaneously, a hyperfunctional glutamatergic state is also recruited during this time (De Witte, 2004).

Typically, this stage is characterized by a dysfunctional hypodopaminergic state and the recruitment of antireward mechanisms, which it may be the source producing negative emotions by engaging activity in the extended amygdala, primarily via the corticotropin-releasing factor, norepinephrine in the hypothalamic-pituitary-adrenal axis and dynorphin (Helga, 2011).

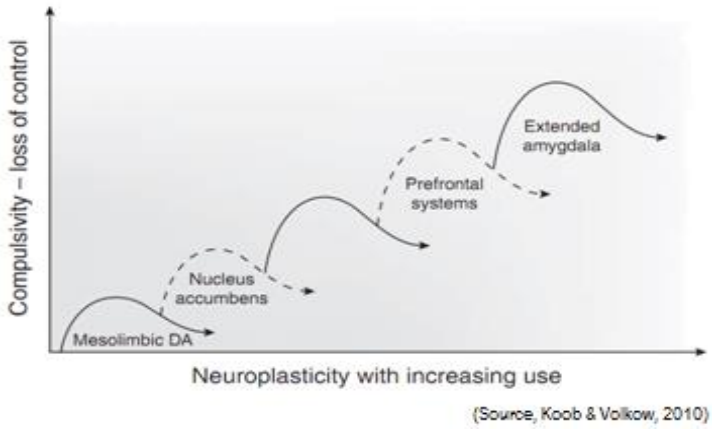


Figure 1.2. **Neuroplasticity with increasing use of drug.**

The schematic figure describes the sequential and cumulative effects of neuroadaptive changes hypothesized to contribute to the neuroplasticity that promotes compulsive drug-seeking (Source, Koob & Volkow, 2010) .

The stage of 'preoccupation/anticipation' (Craving): *a widely distributed network*

The preoccupation/craving stage has been hypothesized to be a key element of relapse which involves a widely distributed network such as the orbitofrontal cortex, dorsal striatum, prefrontal

cortex, basolateral amygdala, hippocampus and insula relating to drug craving and the cingulate gyrus, dorsolateral prefrontal and inferior frontal cortices relating to disrupted inhibitory control (Koob & Volkow, 2010). Generally, the transition to addiction involves neuroplasticity in all of these structures that appears to begin with changes in the mesolimbic dopamine system (Fig. 1.2). The neuroadaptations then gradually relocate from the ventral to dorsal striatum and orbitofrontal cortex, and eventually the process may lead to the dysregulation in a widely distributed network involving the prefrontal cortex, cingulate gyrus, extended amygdala, hippocampus and insula (Fig. 1.2, 1.3; Koob & Volkow, 2010).

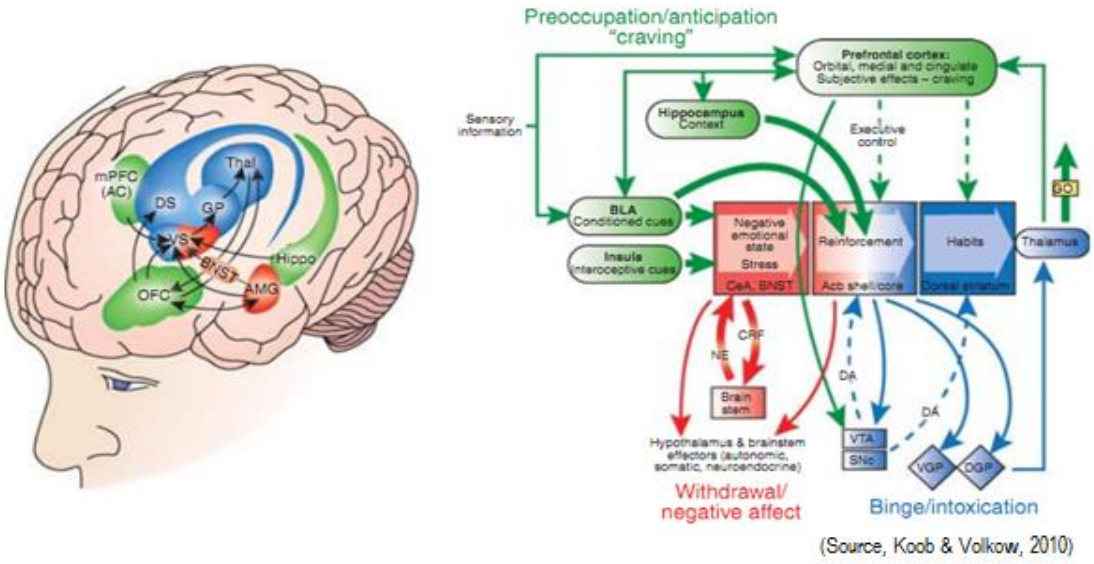


Figure 1.3. **Neural circuits involved with the three stages of the addiction cycle.**

Green/blue arrows, glutamatergic projections; Orange arrows, dopaminergic projections; Pink arrows, GABAergic projections; Acb, nucleus accumbens; BLA, basolateral amygdala; VTA, ventral tegmental area; SNc, substantia nigra pars compacta; VGP, ventral globus pallidus; DGP, dorsal globus pallidus; BNST, bed nucleus of the stria terminalis; CeA, central nucleus of the amygdala; NE, norepinephrine; CRF, corticotropin-releasing factor (Source, Koob & Volkow, 2010).

Pathophysiology of alcohol dependence

The mechanism of alcohol dependence still continues to be studied, but there has been a growing body of evidence from various studies indicating that the mesolimbic dopamine system is the core structure for reward and positive reinforcement (Helga, 2011; Koob & Volkow, 2010; Heinz et al., 2009).

Mesolimbic dopamine system

The chief components of the mesolimbic system are the ventral tegmental area (VTA), ventral striatum including nucleus accumbens (NAc) and their afferent and efferent connections (Fig. 1.4) (Koob & Volkow, 2010).

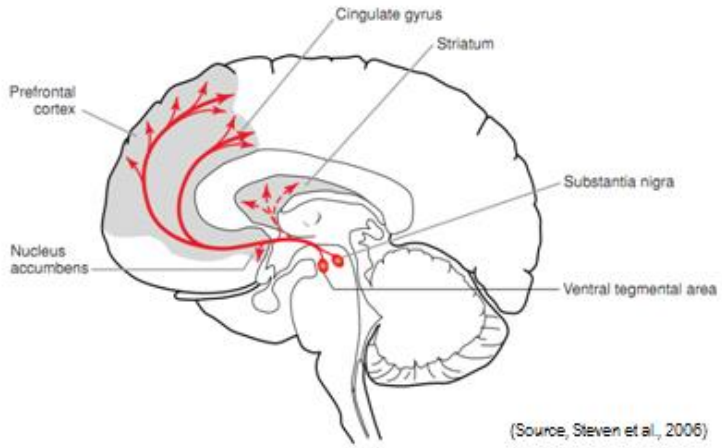


Figure 1.4. **Dopamine projections to the forebrain.**

Projections from the ventral tegmental area to the nucleus accumbens, and prefrontal cerebral cortex, and projections from the substantia nigra to the dorsal striatum (caudate and putamen and related structures) (Source, Steven et al., 2006).

The VTA is situated in the ventral midbrain medial to the substantia nigra and consists of dopamine neurons that project via the medial forebrain bundle to the limbic structures: the NAc, amygdala and hippocampus (called the mesolimbic pathway) and to the medial prefrontal cortex (called the mesocortical pathway) (Fig. 1.4). The NAc, a major component of ventral striatum, consists of two sub-regions which have different morphologies and functions, the shell and the core region. The NAc shell, as part of the extended amygdala, is considered as a limbic structure and engages in drug reinforcement, while the NAc core is a motor region which is more associated with the dorsal striatum (Kelley, 1999). The NAc represents an interface between the limbic neural and motor networks, and may be the important bridge between motivational processes and behavioural action (Doyon et al., 2003), and it is hypothesized that the VTA-NAc is the core region of "brain pleasure centre" mediating the actual pleasure of a reward stimulus as well as reinforcement and motivation for reward-oriented behaviour (Helga, 2011). The source of dopamine to the NAc as well as to the amygdala, hippocampus, and prefrontal cortex (PFC) originates from the VTA of the midbrain (Fig. 1.1 & 1.4) (Steven et al., 2006). In contrast, a significant number of the outward projecting neurons from the NAc are medium spiny GABAergic neurons, and the GABAergic neurons largely connect with the VTA, thalamus, prefrontal cortex and striatum (Kalivas et al., 1993).

The VTA-NAc pathway is regulated by various neurotransmitter systems including the GABA, glutamate, serotonin and acetylcholine systems as well as endogenous opioids and endocannabinoids. All of them influence the reinforcing effects of drugs of abuse, either by acting directly in the NAc or by indirect actions in the VTA (Fig. 1.1; Steven et al., 2006), in which the glutamatergic system, known as an essential excitatory system on the VTA-NAc pathway, plays an crucial role in drug reinforcement and addiction through the control of the mesolimbic dopaminergic pathway. The glutamatergic afferents to the VTA originate from the prefrontal cortex, bed nucleus of the stria terminalis (BNST), laterodorsal tegmental nucleus (LDTg) and lateral hypothalamus. Similarly, the NAc is also innervated by glutamatergic neurons. Most afferents to the NAc core come from the prefrontal cortex and thalamus while the NAc shell receives glutamatergic innervation from the amygdala and hippocampus and prefrontal cortex (Koob & Volkow, 2010). In contrast to the excitatory glutamatergic system, the negative GABAergic feedback system to the VTA regulates the activity of the VTA neurons by providing a modulatory inhibitory tone onto the VTA dopaminergic cell bodies via disinhibition of GABAergic interneurons leading to an inhibition of dopamine release in the NAc (Kalivas et al., 1993). In addition, some other systems such as serotonin, acetylcholine system, and so forth play smaller roles in the VTA-NAc pathway, e.g. the cholinergic afferents that project from LDTg and pedunculopontine tegmental nucleus (PPTg) activate primarily phasic firing of the VTA dopamine neurons via the NAc receptors. Serotonergic projections from raphe nuclei also modulate the mesolimbic dopamine pathways in both the VTA and NAc, and the neuropeptide ghrelin increases dopamine release in the NAc, possibly via a cholinergic mechanism in the VTA (Helga, 2011).

The VTA dopamine neurons can be activated by reinforcers which may be primary stimuli (the actual reward, e.g. addictive substances) as well as conditioned stimuli (e.g. visual or auditory stimuli) (Schultz, 1998), and almost all of them increase levels of synaptic dopamine within the NAc through direct or indirect mechanisms (Wise, 1998). The study results of Doyon and colleagues (2003) on rats showed that a dopamine increase recorded in the NAc was not solely provoked by alcohol (non-conditioned pharmacological effect) but also probably by alcohol-associated cue presentation (conditioned effect). Taken together, this appears to indicate that the VTA-NAc pathway plays a core role in addiction, and stimulation of dopamine release in the

NAc, a core region of the brain reward system, is a crucial property of addictive substances (Wise, 1998; Koob & Volkow, 2010).

Imbalance between reward system and antireward system

Decreased function of brain reward system

Addiction is hypothesized as a cycle of decreased function of the brain reward system and recruitment of the antireward system (Koob & Le Moal, 2008). The taking of acute alcohol results in not only the short-term amelioration of the reward deficit but also suppression of the antireward system (Koob & Le Moal, 2008; Heinz et al., 2009). However, when using long-term administration, the effects of alcohol on the reward system lead to neuroadaptation possibly with synapse plasticity e.g. long-term potentiation (LTP) and long-term depression (LTD) (Anna, 2009), which begins by positive effects on the reward system. Studies on rats showed that alcohol produced a dose-dependent release of dopamine in the NAc, preferentially in the NAc shell when it was given systemically as well as injected locally in the NAc (Di chiara & Imperato, 1998). During this time, a hypodopaminergic state is taken shape by an increase of brain reward threshold and a decrease in the number of dopamine D2 receptors, as a compensatory response with the hyperdopaminergic effects of alcohol on the reward system (Koob & Le Moal, 2008). Imaging studies in drug-addicted humans have consistently shown long-lasting decreases in the numbers of dopamine D2 receptors in drug abusers compared with controls (Volkow et al., 2003; Heinz et al., 2004).

Recruitment of antireward system

Simultaneously, an opponent system, known as antireward system, also causes the neuroadaptation, but in the opposite direction, such as up-regulation of NMDA receptors (N-methyl-D-aspartate receptor) which may originate from the effects of alcohol on the glutamatergic neurotransmission. Alcohol stimulates $GABA_A$ receptors and inhibits the function of glutamatergic NMDA-receptors (Kalivas & Volkow, 2005; Beck et al., 2011). Such effects in the long-term lead to the reduction of effects of glutamate on NMDA receptors and thereby result in compensatory up-regulation of NMDA receptors (Heinz et al., 2009). The antagonistic adjustment of the antireward system tries to achieve a balance between the two systems, also known as allostatic state. The allostasis is defined as stability through change. Allostasis is quite

more complex than homeostasis and has several special characteristics that differ from homeostasis (Sterling & Eyer, 1988 cited by Koob & Le Moal, 2008). Allostasis involves a feed-forward mechanism which is rather different from the negative feedback mechanisms of homeostasis. For instance, when an increased need produces a signal in homeostasis, negative feedback mechanism is started to correct the need to keep it at a constant level. In contrast, in allostasis, there is continuous re-evaluation of need and continuous readjustment of all parameters toward new set points (Koob & Le Moal, 2008). Also, when an alcohol-dependent patient abstains from alcohol, a new imbalance turns up due to the loss of effects of alcohol on the system. At that time, this condition discloses the hypodopaminergic as well as hyperglutamatergic state which originates from its effects on the system over a long period of time (Fig. 1.5). Microdialysis studies on rats show that ethanol withdrawal is associated with increases in glutamate in the striatum, nucleus accumbens and hippocampus approximately 5–8 hours after cessation of ethanol inhalation, with a maximal value at 12 hours (Rossetti & Carboni, 1995; Dahchour & De Witte, 1998). Then, the body can be on impulse for a change to achieve a new balance, a new allostasis, although it is likely that the new balance may not be healthy, but it is "appropriate" to environmental demands (Koob & Le Moal, 2008). Alcohol dependence thus can be viewed as a dynamic phenomenon represented by a transition from neuroadaptation to pathophysiology (Clapp et al., 2008; Koob & Le Moal, 2008).

Motivation of compulsive alcohol seeking

Based on the fact that the brain is a network of systems working in equilibrium (De Witte, 2004; Becker, 2008), the imbalance may be just what motivates alcohol-dependent patients after abstinence to compulsively seek alcohol with the goal of restoring the balance which the patients had stabilized and adapted to during a long period of alcohol consumption before abstinence (Koob & Le Moal, 2008). The requirement of restoring the balance lasts a short or long time, depending on the time it takes to re-establish a new balance which is contingent on many factors e.g. addictive level of patient, environmental factors, willpower of patient, genetic variables, etc. (Christopher, 2006; Koob & Le Moal, 2008). Evidence reflecting indirectly the progression can be found in a follow-up study of alcohol dependence of Heinz et al. (1996) indicating that down-regulation of dopamine D2 receptor in the ventral striatum is almost prominent just after detoxification and recovers during abstinence. This result appears to suggest that there is

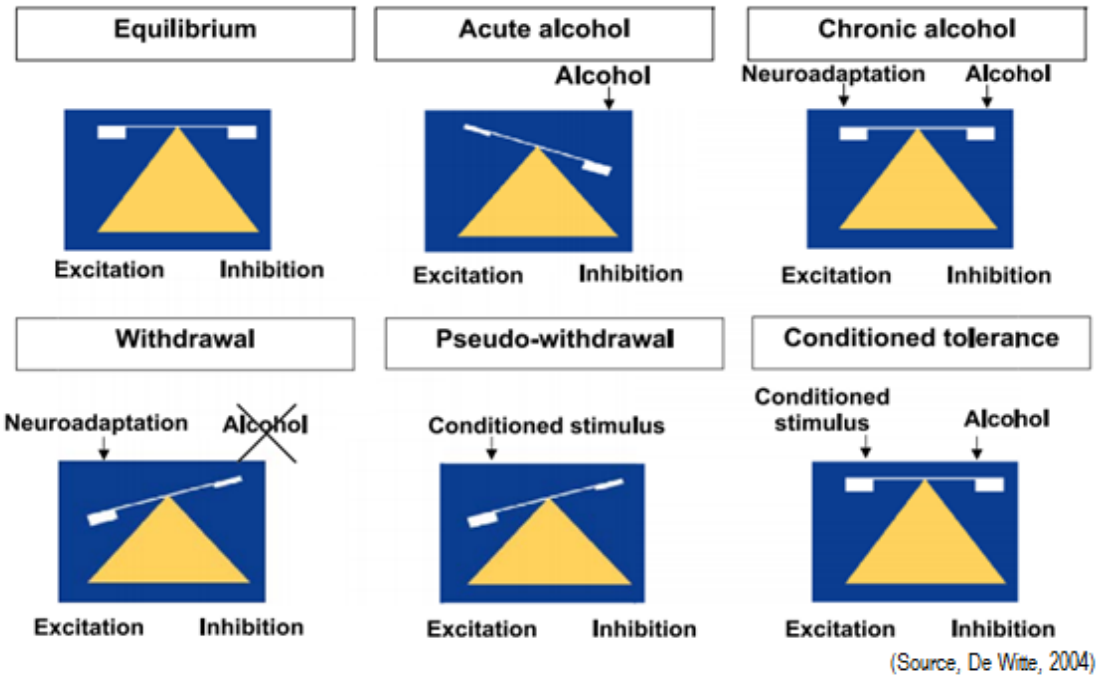


Figure 1.5. **This figure illustrates the brain (triangle) that is controlled by different excitation and inhibition processes to maintain the brain in a regular equilibrium.** Acute alcohol disrupts the equilibrium by enhancing the inhibitory processes (mainly GABA and taurine) that indirectly increase dopamine release via inhibiting $GABA_A$ interneurons in the VTA-NAc. Chronic alcohol consumption causes neuroadaptation (up-regulation of glutamate) to counteract the inhibitory action of alcohol. Withdrawal of alcohol results in an overexcitation state of the brain due to the excess of neuroadaptative excitatory processes. Conditioned stimulus alone may lead the brain to a state similar to withdrawal state called mini-withdrawal. Conditioned tolerance may also occur through the presence of alcohol together with conditioned stimulus (Source, De Witte, 2004).

neuroadaptation in the reward system after alcohol withdrawal in order to re-establish the balance, and the process moves towards complementing the hypodopaminergic state. Therefore, the slow or fast recovery of central dopaminergic neurotransmission can be a sign to predict the probability of either relapse or recovery among detoxified alcoholics (Heinz et al., 1996, 2004).

Role of alcohol-associated cues in alcohol dependence

Alcohol-associated cues as conditioned stimuli

One of the characteristics formed during alcohol dependence, which plays an important role in relapse mentioned in a series of previous studies, is cue-related response (Schultz, 1998; Wise, 1998; Drummond, 2000; Doyon et al., 2003). The cues can serve as conditioned stimuli that can

encourage alcohol consumption (O'Brien et al., 1998; Drobes, 2002). Alcohol and other addictive substances act as 'instrumental reinforcers', which increase the power of responses that produce them, leading to drug self-administration or 'drug taking'. Environmental stimuli such as time, space, pictures, and so on that are closely associated with the effects of self-administered drugs obtain incentive salience through the process of Pavlovian conditioning (Everitt & Robins, 2005). The underlying activation of neural structures involved in maintaining the incentive salience state makes addicts vulnerable to long-term relapse. The way of response to these stimuli is presumably stored as alterations in synaptic weights and, eventually, after a long time, by physical remodelling of synaptic connections (Berke & Hyman, 2000). In previous imaging studies (Braus et al., 2001; Wrase et al., 2007; Park et al., 2007; Beck et al., 2009; Heinz et al., 2009), such alterations appear to be evidenced by a significantly difference in activation in brain regions involving the mesolimbic system, especially the ventral striatum including the NAc, in alcohol-dependent patients compared with healthy controls when elicited by alcohol-associated cues.

Enhanced sensitivity to the cues

A hypodopaminergic state is exposed during early detoxification and abstinence possibly due to the lack of effects of alcohol on the reward system. Studies on rats following alcohol self-administration training showed that when they self-administered alcohol, a concurrent rise in dopamine levels was produced in the NAc, whereas a withdrawal from alcohol decreased dopamine release in the NAc (Diana et al. 1993; Weiss et al., 1993; Rossetti et al., 1992). Concurrently, a hyper-antireward state also breaks out due to the loss of the factor inhibiting the antireward system. This phenomenon is illustrated in the Fig. 1.5, where the loss of alcohol-associated inhibition on the glutamatergic system (especially NMDA receptors) may result in hyperexcitation and clinically manifest as withdrawal symptoms (Spanagel, 2003; De Witte, 2004). Hence, it seems that the imbalance between the two systems is the source leading to enhanced sensitivity to the conditioned stimuli with the goal of compensating deficiency of alcohol or addictive substances in order to balance the systems (Koob and Volkow, 2010). For instance, a study of McClernon et al. (2009) on the effects of withdrawal on cue reactivity indicated that abstinence from smoking can dramatically potentiate neural responses to smoking-related cues in the brain regions which are in charge of visual sensory processing, attention and

action planning. Besides withdrawal, other factors e.g. acute intoxication, family history, gender, expectancy or drug availability, genotype also show their influences on response sensitivity to the cues (Filbey et al., 2011). A small, priming dose of alcohol, for example, enhanced the effect of olfactory cues in the NAc, medial frontal, orbitofrontal and posterior cingulate cortex recorded in a study by Bragulat et al. (2008).

Transition in response to the cues

As addiction progresses from initial drug use to a dependence syndrome, the neurocircuitry and neurochemistry shift from a behavioral system based on dopamine release in the NAc with acute administration (signaling initial reward and beginning the process of conditioned learning) to a behavioral system predominantly based on glutamate (initiating the process of drug reinstatement or relapse) (Ross & Peselow, 2000). Therefore, the imbalance after withdrawal accompanied with the excessive activity of glutamate indicates that the glutamatergic pathways from the prefrontal cortex, amygdala and hippocampus to the NAc and VTA play a major role in triggering relapse (Fig. 1.3) (Kalivas et al., 2005; Heinz et al., 2009; Koob and Volkow, 2010). Furthermore, in the way of response to alcohol-associated cues, cue-induced activation of the anterior cingulate and adjacent medial prefrontal cortex involving the ventral striatum may mediate an attention response to alcohol-associated cues while cue-induced dopamine release in the dorsal striatum can trigger relapse into drug-taking behaviour (Ito et al., 2002; Heinz et al., 2004). Robbins and Everitt (2005) have proposed that the initial reinforcing effects of drugs of abuse may activate the ventral striatum, but when the drug taking transitions into habitual drug-seeking behaviours, activation of the more dorsal striatal regions predominates. The dorsal striatum does not appear to have a major role in the acute reinforcing effects of drugs of abuse but appears to be recruited during the development of compulsive drug seeking (Everitt & Robbins, 2005). This implies that the dorsal striatum is crucial for habit learning, e.g. for the learning of automated responses, and may thus contribute to the compulsive character of dependent behaviour. In other words, in addicted individuals, cue-elicited craving tends to preferentially elicit dopamine release in more dorsal striatal structures, which is thought to reflect a transition from a ventral striatal reward-driven phenomenon to a dorsal striatal stimulus-response habit formation (Berke & Hyman, 2000), in which reward plays a lesser role. For this reason, it is likely that habit expressed by dorsal striatum activation can play an important role in

forming a fast, easy and automatic response relating to alcohol-associated cues. In other words, the characteristics of activation of this structure to specific stimuli can be referred in order to predict the addictive level of a patient. The hypothesis is supported by the study result of Vollstädt-Klein and colleagues (2010) indicating that the dorsal striatum of heavy drinkers was activated more strongly than that of light drinkers, whereas light social drinkers showed stronger cue-induced fMRI activations in the ventral striatum and prefrontal areas than those of heavy social drinkers.

In summary, it appears that alcohol dependence is a dynamic process in which there is transition to-and-fro between the stages of addiction. Furthermore, the response features to alcohol-associated cues can reflect the stages of the disorder whereby we can predict the alcohol dependent status of a patient. In other words, the reactivity of the brain circuits to alcohol-associated stimuli may serve as a biomarker to help predict relapse as well as treatment efficacy (Koob and Volkow, 2010).

fMRI AND CLASSIFICATION TECHNIQUES

fMRI data

Functional magnetic resonance imaging (fMRI) is an advanced non-invasive medical imaging technique that can give high quality visualization of brain activation through changes in blood flow or oxygenation resulting from sensory stimulation or cognitive function (Ogawa et al., 1990). It therefore has been often used in studies of brain function e.g. to investigate how the healthy brain functions, how it is affected by different diseases, how it attempts to recover after damage and how drugs can modulate activity or post-damage recovery, etc.

fMRI experiment

During the course of an fMRI experiment, a series of three-dimensional images of a subject's brain activity are recorded while he is performing a set of tasks, known as fMRI paradigm. Then, the images from different subjects are analyzed to detect differences of brain activation in the brain regions of interest between the investigated groups of subjects. Therefore, designing an appropriate paradigm is one of the most important tasks for an fMRI experiment. Currently, there

are two commonly used approaches, “block design” and “event-related design” (Edson & Gareth, 2006).

The block design is the simplest approach. The different experimental conditions are separated into extended time intervals, or blocks. The cycles of periods of task and rest (conditions) are arranged alternately. This design allows maximization of signal-noise ratio (SNR) but also has some disadvantages. Repeating the same task may lead to the subject anticipating the task and sometimes even the response. This may considerably confound the results.

The event-related design is a more flexible and complex approach. The order of the stimuli is often randomized and even the time between stimulus presentations also varies (interstimulus interval) to prevent anticipation of the task. However, the disadvantage of this design is the low SNR. This is due to the fact that the task state is not sustained for long periods, leading to a less intense vascular response (Edson & Gareth, 2006; Graeme et. al., 2008).

Apart from the task-driven fMRI just described, recently interest has been growing in the application of the technique at rest, termed resting-state fMRI (RS-fMRI). The RS-fMRI is applied to evaluate synchronous activations between brain regions that take place in the absence of an explicit task or stimulus. Although this is a relatively new method, it has shown promise in providing diagnostic and prognostic information for neuropsychiatric disorders (Lee et al., 2012).

fMRI scanner

The MRI scanner creates a powerful magnetic field (0.2 - 3T), which causes some nuclei (predominantly hydrogen nuclei or protons in the water) in our body to align parallel or anti-parallel to the applied magnetic field, according to their spin. Pulses of radio frequency (RF) then are applied to excite the protons (90° excitation RF pulse) and systematically flip the spins of the aligned protons. Since the application of RF pulse disturbs the spin system in the strong static magnetic field, there is subsequently a process to return to equilibrium (pre-excited stable state) when the RF is turned off. This relates to exchange of energy between the spin system and its surroundings, and as the protons return to the lower energy state, radio waves are emitted. They are then recorded and processed to construct an image of the scanned area. The protons can return to the stable state only by dissipating their excess energy to their surroundings. The process is called spin-lattice relaxation, T1 relaxation. The rate of restoring the equilibrium is

characterized by the spin-lattice or longitudinal relaxation time, T1. Nonetheless, the spins exchange energy not only with the surrounding lattice but also among themselves. The process is known as spin-spin relaxation or transverse relaxation, T2 relaxation. In this relaxation, the spins do not dissipate energy to their surroundings but instead exchange energy with each other. The process generally takes place faster than the spin-lattice relaxation. In order to improve recorded image quality, a technique of spin echo sequences is used by the application of an 180° refocusing pulse to eliminate the effects of static field inhomogeneities. The tradeoff of this technique is a fairly long scan time. The T2* imaging used in fMRI does not use this refocusing technique, so the resolution of the images is reduced (to approximately 3 mm), but the sensitivity to the relaxation processes is increased. Besides, with the system equipped with echo planar options, the image acquisition interval is very short, typically every 0.5-4 seconds for each scan (Clare, 1997; Mathews, 2001; Weishaupt et al., 2008).

Image contrast between gray matter, white matter and cerebrospinal fluid can be optimized by appropriately weighting the relaxation times. For example, T1-weighted images provide clear contrast between gray matter and white matter, so they are often used to create high-resolution (approximately 1 mm) 3D structural images taken in slices at a single point in time. In contrast, thanks to the advantage of very short acquisition time, T2*-weighted images are employed to analyze brain activity under impact of specific stimulation (Mathews, 2001; Weishaupt et al., 2008; Yang et al., 2011).

fMRI signal (BOLD signal)

Brain activity is indirectly recorded via the blood oxygenation level dependent (BOLD) signal. The application is based on the paramagnetic property of deoxygenated haemoglobin. Normal blood can be seen simply as a concentrated solution of haemoglobin (10-15 gm haemoglobin/100 cm^3). When haemoglobin is attached to oxygen (oxygenated haemoglobin), it becomes diamagnetic, while deoxygenated haemoglobin is paramagnetic (Pauling and Coryell, 1936 cited by Mathews, 2001). Paramagnetic materials are attracted by the applied magnetic field, i.e. they strengthen the magnetic field. They therefore increase the T2* relaxation rates (i.e. decreases T2 time). This attenuates the T2* magnetic resonance signal. In contrast, diamagnetic materials are repelled by the applied magnetic field, so they increase the signal. In other words, a change in haemoglobin oxygenation induces a corresponding change in the recorded signal intensity. This

characteristic is exploited in the investigation of task-induced neuronal activity due to the coupling of hemodynamic response to neuronal activation. A locally increased blood flow and volume in the brain region which becomes active appears to be a consequence of increased energy utilization at the synapse e.g. a local increase in glucose and oxygen consumption (Mathews, 2001; Logothetis & Pfeuffer, 2004). However, the increase in the blood supply exceeds the metabolic needs, which leads to increased blood oxygen concentration in the activated region (Fox & Raichle, 1986). As a result, the increase of blood oxygenation increases the T2* signal recorded in the region. This is the basis for the BOLD fMRI. Thus, the BOLD signal is a secondary effect of neuronal activation, and there is the time delay in the hemodynamic response, the peak of which occurs 4-6 seconds after the neuronal activity (Mathews, 2001). In other words, the recorded fMRI is the image indirectly reflecting neuronal activation through hemodynamic response. Accordingly, in fMRI analysis a hemodynamic response function (HRF) or impulse response function is often incorporated in the computational models of neuronal activation by convolving with the neuronal response evoked by the stimulus that has been designed in fMRI paradigm (stimulus function) in order to give a hemodynamic response (Friston et al., 2007).

fMRI image

A typical 20-minute fMRI experiment produces a series of 3D brain images (volumes or scans), each of which contains approximately 170,000 voxels (e.g. for an image matrix of 64 x 64 x 42). First, data are collected from an fMRI scanner on the subject undergoing an experiment designed to activate the neuronal responses in the brain regions of interest. The recorded intensity values of BOLD signal are processed and then normalized to range between zero and a fixed constant e.g. between 0 and 1500. The time taken to acquire a single fMRI image (volume) is of the order of several seconds. Thus, each 2D plane (slide) of the 3D fMRI image (volume) records brain activity from different points in time (Burge, 2007); and each volume is stored in a chronological record in a three-dimensional matrix [x, y, z], the elements of which store image resolution (pixel or picture element) representing the intensity of activation. For instance, in a 3D brain image matrix with dimensions of 64 x 64 x 42, there are 42 slides. Each slide is a two-dimensional matrix of 64 columns and 64 rows comprising 64 x 64 elements, known as voxels, that store image resolution values (voxel attribute) between 0 and 1500 and

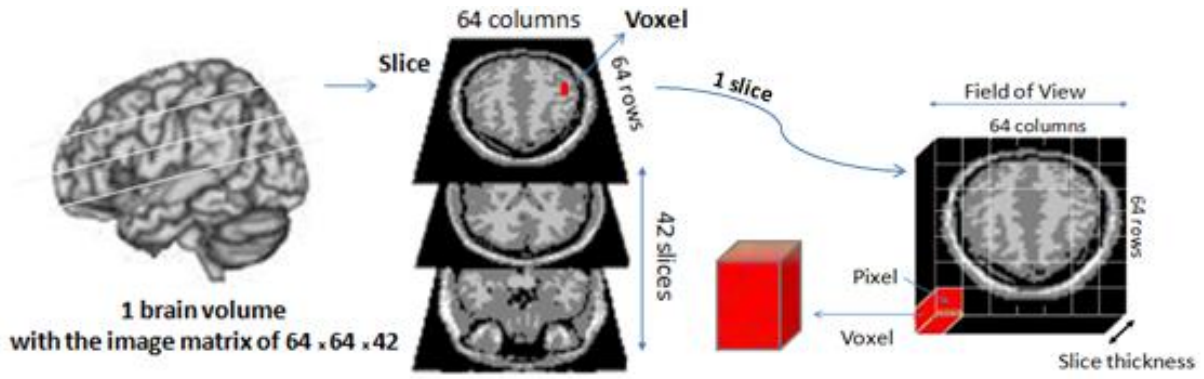


Figure 1.6. **Illustration for a volume of 3D brain image with dimensions of 64 x 64 x 42.**

that represent cubes of tissue about 2-4 millimeters (volumetric pixels) depending on slice thickness, field of view (FOV) and size of image matrix (Fig. 1.6).

Pre-processing raw fMRI data before analysis

fMRI data are susceptible to a large number of artifacts which can roughly be divided into scanner-induced artifacts e.g. radio frequency, gradient artifact, etc. and physiological artifacts e.g. motion, respiration, heartbeat, contamination from large veins and arteries in the brain, etc. (Graeme et. al., 2008; Lindquist, 2008). Consequently, to minimize non-task-related variability in the recorded image data within-subject as well as between-subject for validity of statistical assumptions before the data are analyzed, they need to be pre-processed. The pre-processing comprises a series of steps that can be roughly divided into anatomical and functional steps.

The functional steps include temporal and spatial processing. For temporal processing termed slice timing correction, each slice in each volume is acquired at slightly different points in time. Therefore, it is necessary to adjust the data so that it appears as if all voxels within one volume had been acquired at exactly the same time. Spatial processing is designed to remove movement effects termed motion correction or spatial realignment. Besides, spatial and temporal smoothing with a Gaussian kernel is often performed to improve the SNR of imaging data and to reduce differences between the activation patterns of subjects (Etzel et al., 2009).

The anatomical steps include spatial coregistration and normalization. Since fMRI is typically of low spatial resolution and provides relatively little anatomical detail, the coregistration is

designed to eliminate differences between structural and functional images in order to be able to map the results obtained from functional data onto a structural MR image with high resolution for presentation purposes. When performing group analysis to make population inferences, all individual brain images recorded for all subjects are assumed to be registered so that each voxel is located in the same anatomical region. For this, spatial normalization is applied to register all of the fMRI images into the same standard space e.g. Montreal Neurological Institute (MNI) or Talairach space. Without the preprocessing prior to analysis, the result of statistical analysis would be invalid (Lindquist, 2008).

fMRI analysis

Since fMRI was invented in the early 90s, it has become one of the widely used non- invasive techniques for investigating human brain activity. Along with its development, the analysis methods of fMRI data have appeared. Today, fMRI analysis has been used for three main applications including localization of brain activation, connectivity and classification/prediction (Fig. 1.7).

Localization of brain activation

Individual-voxel-based approach

A few years after fMRI was invented, the traditional analysis methods of approach to fMRI came into sight and put into application. This approach has focused on characterizing the relationship between cognitive variables and individual brain voxels. In other words, the fMRI analysis to indicate the activated brain regions by specific tasks is performed separately at each voxel (mass-univariate approach). The analysis uses statistical regression and hypothesis testing based upon the general linear model (GLM) or

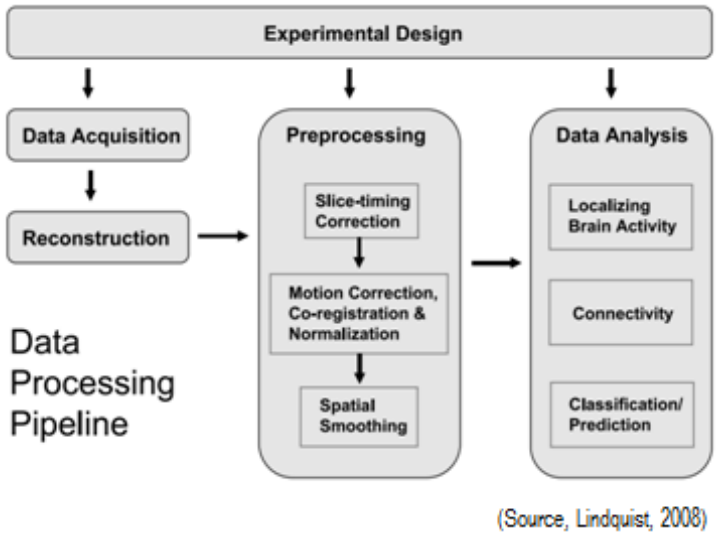


Figure 1.7. **The fMRI data processing pipeline illustrates the different steps involved in a standard fMRI experiment** (Source, Lindquist, 2008).

discriminant analysis techniques such as multivariate regression to test hypotheses about regionally specific effects. These techniques commonly make linearity and Gaussian noise assumptions and eliminate the time factor in inference (Friston et al., 2007; Lindquist, 2008). Due to limitations of the statistical approach, about ten years later, an alternative approach has appeared. The new approach emerged from the Bayesian theory. In contrast to statistical inferences about the data, given the effect is zero, Bayesian inferences are based on conditional inferences about an effect, given the data (Friston et al., 2007).

Multivariate approach

Although individual-voxel-based methods are still being used widely due to their simplicity (Friston et al., 2007), they have exposed the limits of what can be learned about cognitive states by analyzing voxels individually (Normand et al., 2006). The limits have promoted the development of new approaches where the fMRI analysis considers groups of voxels rather than a single voxel. The analyses range from the "searchlight" approach which multivariately examines the information in small groups of voxels centered on each voxel in the brain (Kriegeskorte et al., 2006) to multivoxel pattern analysis which can detect patterns across voxels in fMRI data (Mourao-Miranda et al., 2005; Normand et al., 2006; Etzel et al., 2009). The primary advantage of these methods over individual-voxel-based methods is increased sensitivity (Normand et al., 2006).

Connectivity

The brain is the center of the nervous system and is made up of nerve cells (neurons). Its function is to exert centralized control over the other organs of the body. To take on this responsibility, single neurons do not work independently but rather function in large aggregates (neuronal groups), known as functionally specialized brain regions e.g. motor areas, sensory areas, visual cortical areas, etc. (Mathews, 2001; Bear et al., 2007). Furthermore, between the different functional regions there are also connections or interactions, and when responding to a specific stimulus, several relevant brain regions would be activated interactively (Bear et al., 2007). In other words, neurons within the brain regions as well as between these regions that are in charge of this response have high interactions (correlations). Due to the coupling of neural activation and local haemodynamic response characterized by voxel attribute (BOLD signal),

there exist correspondingly high correlations among relevant voxels. For examples, the time-series of BOLD signal recorded on one voxel looks correlatively like the time-series of BOLD signal on another voxel. If they are located adjacent to each other (e.g. in the same brain area), they are called as a cluster. If they are located far away from each other e.g. one is in parietal cortex and the other is in frontal cortex, they are thought to be connected to each other somehow. However, correlation doesn't imply direction. Hence, in 1993, Friston introduced the two different approaches to investigating connectivity in functional neuroimaging.

Functional connectivity

The first approach is defined as functional connectivity. This approach is focused on pairwise interactions often in terms of correlations or covariances between voxels or brain regions of interest. It does not provide any direct insight into how these correlations are mediated (undirected association) (Friston, 1994; Lindquist, 2008). The simplest method of the approach is to compare correlations between brain regions of interest, or between a "seed" region and the other regions or voxels throughout the brain (regional correlation). However, it becomes problematic when the number of correlations grows because it needs to correct for multiple comparisons, and it is difficult to summarize the patterns of correlation. Alternative approaches use multivariate methods e.g. principal components analysis (PCA) and independent components analysis (ICA), etc. to detect task-related patterns of brain activation without making any a priori assumptions about its model (Lindquist, 2008).

Effective connectivity

The second approach is defined as effective connectivity that shows the directed influence of one brain region on the others. The approach incorporates additional information e.g. anatomical connections into the analysis. In addition, a simultaneous interaction of several neural elements is also considered to explicitly measure the effect of one element on the other (Friston, 1994; Lindquist, 2008). In regard to measurement methods of effective connectivity, Büchel & Friston (1997) introduced structural equation modeling (SEM), also known as path analysis, which is used to investigate significant changes in the relationship between neural systems in the dorsal visual stream caused by shifts of attention. Hojen-Sorensen, Hansen, & Rasmussen (2000) used another approach based on Bayesian network theory such as Hidden Markov Models

(HMMs) to learn a model of activity within the visual cortex from visual stimuli. Recently, dynamic causal models (DCMs) have been introduced by Friston and colleagues (2003) with the goal of modeling effective neural connectivity (Friston et al., 2007). The technique is also a branch of dynamic Bayesian networks, and it is applied to characterize brain activity at the level of neural networks and their dynamics. Today, the technique has become an important tool of neuroimaging analysis and has an important impact on the development of theoretical neurobiology and clinical biomarkers (Seghier, 2010).

Classification/prediction

Another application direction relating to our research is classification of fMRI, also known as pattern recognition. fMRI classification is a technique of separating fMRI data into different classes, i.e. providing a criterion for determining whether the BOLD response of a subject at a particular time during the experiment characterizes a specific cognitive state, a neuropsychiatric disorder or not (Ye Yang, 2010). The specific tasks for a study of fMRI classification is to construct patterns from fMRI images, to build up a classifier from the labelled patterns of training data and then to test the classifier on the unlabeled and unseen patterns of testing data (i.e. to use the classifier to label the unseen patterns of testing data) (Pereira et al., 2009).

Constructing patterns from fMRIs

This is the step of constructing features for a pattern from an fMRI (feature construction). If all voxels of an fMRI image are used as features of the pattern, the pattern contains very large number of features (e.g. approximately 170,000 features for a 3D image matrix of 64 x 64 x 42). For a set of patterns with such very large size of feature, classification performance of the patterns can be reduced significantly (Pereira et al., 2009). Further, for a brain response to a given specific stimulus, not all of the voxels are activated significantly (Etzel et al., 2009). This implies that there may be uninformative voxels in a classification. Hence, methods to reduce the number of features for a pattern extracted from an fMRI have been developed. They are divided into two main approaches.

The first approach is to select informative voxels (features) from an fMRI (feature selection). There are two methods for this approach: scoring/filtering and wrapper method. The first involves ranking the features based on a given criterion and selecting the best in the ranking. The

latter involves performance of a classifier. For this, firstly, all the features are considered and then they are removed gradually while the performance increases. The method is known as recursive feature elimination, and it can be accomplished by repeatedly training and applying the classifier in cross-validation within the training set (Pereira et al., 2009).

The second approach is to reduce the number of features of an analyzed pattern (feature reduction or dimensionality reduction). The approach focuses on correlation between features. The commonly used methods in this approach may be named as singular value decomposition/principal component analysis (SVD/PCA), independent components analysis (ICA), etc. The general nature of these methods is that they transform the original feature space into a new, low-dimensional feature space. This yields a new dataset matrix with a reduced number of features. In addition, another well-known approach is to use a combination of the two approaches (Lindquist, 2008; Pereira et al., 2009).

Previous works of fMRI classification

A classifier is a function that takes features of a pattern to predict its label. The classifier is formed from learning characteristics of labelled patterns of a training dataset. Such an approach is known as machine learning. Along the time line, the development of technology and solutions of pattern recognition applied in fMRI classification is still progressing. Several prominent milestones of the development may be mentioned. In 1936, Fisher introduced linear discriminant analysis (LDA) that computes a hyperplane in the input space so that it maximizes the ratio of between-class variance to within-class variance (Fisher, 1936). The method can work well with linear data (Rätsch, 2005). However, it is not sufficient for fMRI classification, where data sometimes are not linearly separable. In the late 60s, Cover and Hart (1967) introduced k-Nearest neighbour classification. Here, the k points of training data closest to the test point are identified, and a label is assigned to the test point by a majority vote between the k points. This method is simple, but it requires expensive computation and a large memory to store the training data. Turing (1992) first proposed artificial neural network for classification. Afterwards, the technique has become one of the commonly used approaches for classification. Also in the 90s, a statistical learning theory appeared (Boser et al., 1992; Vapnik, 1998; Vapnik, 2000), which provided conditions and guarantees for good generalization of learning algorithms. Recently, large margin classification techniques have emerged as a practical result of the theory of

generalization. The two large margin classifiers frequently mentioned are support vector machines (SVMs) (Boser et al., 1992, Cortes and Vapnik, 1995) and boosting (Valiant, 1984). These methods have demonstrated highly competitive performance in many studies of fMRI classification reported (Etzel et al., 2009). In 2003, Mitchell et al. (2004) introduced a Gaussian naïve Bayesian network which was used to classify instantaneous cognitive states of a subject while reading a book or looking at a picture. Burger et al. (2007) applied dynamic Bayesian network, a data-driven modeling technique, to identify functional correlations among regions of interest with the goal of classifying healthy and dementia fMRI data.

Research problems

Although encouraging achievements have been reported in the studies of fMRI classification with predictive accuracies between 70 and 90% (Shinkareva et al., 2006; Demirci et al., 2008a; Demirci et al., 2008b; Takayanagi et al., 2011), they are usually difficult to be generalized with larger data sets (Demirci et al., 2008b). Several reasons have been mentioned such as limited number of subjects investigated, bias in classification, variability between operators, scanning equipment and parameters, and variability between subjects and between different times of measurement even within the same subject (Demirci et al., 2008b). This indicates the complexity of fMRI data as well as the unstable reliability of classification decision achieved from machine inference, whereas a classification decision for each individual patient requires very high accuracy and reliability.

For these reasons, while waiting for the technological solutions to meet our demands in clinical practice, it is necessary to find an alternative solution of fMRI classification which can help us avoid complete dependence on machine inference. This can be realized if we can check the compatibility between the classification decision for a pattern obtained from machine and its activation image. In other words, a thorough understanding of the classified pattern and of the classification decision for the pattern obtained from machine may bring the solution to light, and it may be a feasible approach to realizing diagnostic functional imaging of neuropsychiatric disorders in clinical practice.

For alcohol dependence, several lines of evidence have shown significant differences in response to alcohol-associated cues between detoxified alcoholics and healthy controls (Braus et al., 2001;

Myrick et al., 2004; Wrase et al., 2007; Park et al., 2007; Beck et al., 2009; Heinz et al., 2004, 2009) and between relapsers and abstainers (Schacht et al., 2011; Beck et al., 2012) based on fMRI measurement. This indicates that the fMRIs hold important information of differences among the investigated groups or in other words, the fMRIs can be used as useful biomarkers for diagnosis as well as prognosis in alcohol dependence. However, the results of these studies have obtained from a statistical analysis between the different groups. Such an analysis is designed to identify the brain regions showing significant differences in response to the given stimuli between the two investigated groups (difference between groups) rather than to provide observations of the differences between individual subjects of the two groups to be used for classification (difference of individuals) (Demirci et al., 2008b; Lindquist, 2008; Van Horn & Poldrack, 2009; Farah & Gillihan, 2012). These problems have motivated us to conduct the dissertation.

AIMS

The overall objective of this dissertation is to develop a framework which enables the identification of alcohol dependence as well as prediction of relapse risk in clinical practice using fMRI. The specific objectives were focused as follows:

(1) To design and validate a classification algorithm for diagnosis and relapse prediction using fMRI in such a way that the classification results are interpretable
(2) To approach imaging based on the findings gained from the classification algorithm for the investigated fMRI data
(3) To validate the approach

METHODOLOGY

Outline of the whole approach

The approach was designed as a means of converting the findings of machine-based classification into our understanding of classification rules on functional imaging. For this, firstly, classifiers were formed from given classification algorithm and used as intermediate exploratory instruments, instead of us seeking the rules of recognizing the investigated patterns (characteristics for recognition). Then, based on the findings as well as working rules of the

classifiers, the rules for diagnostic functional imaging in clinical practice can be uncovered (Fig. 1.8).

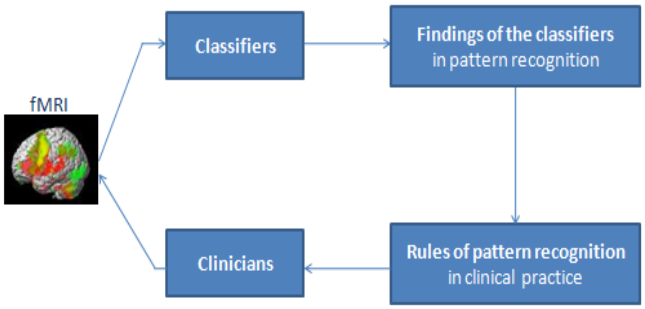


Figure 1.8. **The framework for the approach.**

Based on the idea, we partitioned the whole approach into smaller approach steps starting from classification algorithm to imaging approach. From this point, the first studies were formed based on classification algorithms. Then, these machine-based classification algorithms would be replaced with a diagnostic imaging approach in the next studies. Correspondingly, the algorithms for the studies have been changed continuously and appropriately to expose the whole approach that can lead us to realize diagnostic functional imaging in practice. Hence, the algorithm for the whole idea is not a single algorithm rather than it is just a synthesis of the whole approach. In other words, we would like to build a framework for this approach (Fig. 1.8). For this reason, each study was conducted using a different methodology for its specific objective. To facilitate the presentation, we arranged the methodologies, results and discussions of the studies in separate chapters.

Specifically, the first study was to demonstrate feasibility of splitting observation on the whole brain into multiple observations on multiple relevant brain regions involved in alcohol dependence using fMRI (chapter II). The second study was to demonstrate the validity of predictive inference based on multiple lines of evidence collected from several brain regions of interest in relapse prediction using fMRI (chapter III). These two studies served for specifying the algorithm and important brain regions involved in alcohol dependence in fMRI classification. The third study was to offer an imaging approach based on the findings of the first and second studies (chapter IV). Finally, we introduced two feasible applications of the approach in clinical practice (chapter V).

CHAPTER II

FORMATION OF FUNCTIONAL ROIs IN fMRI CLASSIFICATION

Introduction

To make an imaging diagnosis in clinical practice we cannot observe and analyze the whole brain but rather we should focus this observation on several relevant brain regions because with such an approach the observed results are easier to identify and interpret. However, whether or not and how the approach is feasible for fMRI classification in alcohol dependence remains the question that we would like to clear up in the first study. Specifically, this study was to look for the clues with which we could identify the appropriate way of feature selection for each brain region of interest that can yield a high performance of classification in alcohol dependence.

Materials and Methods

Materials

Participants

Fifty alcohol dependent patients diagnosed according to ICD-10 and DSM-IV criteria and 57 healthy subjects were recruited for the study. All participants were right-handed volunteers who accepted participation after the research procedures had been fully explained to them. The study was approved by the Ethics Committee of Charité Universitätsmedizin Berlin, Campus Mitte in Berlin in Germany. All the participants were over 18 years of age, ranging from 22 to 69 years (mean = 41.8; standard deviation = 12.1; 81 males and 26 females). In addition, the subjects had no other psychiatric axis I disorders, no past history of dependency or current abuse of other drugs, which was verified by random urine drug testing and interviews. Before the fMRI experiment, the patients had to be abstinent from alcohol for at least 7 days in an inpatient detoxification treatment program.

Data acquisition

The data were acquired with a 3 Tesla scanner (Siemens, Erlangen, Germany). The imaging sequence was an ascending T2*-weighted echo planar sequence with 42 axial slices (repetition

time (TR) = 2.41 s, echo time (TE) = 25 ms, flip angle = 80^0, field of view (FOV) = 192 x 192 mm^2, slice thickness = 2 mm, gap between slices = 3 mm, acquisition matrix = 64 x 64 and voxel size = 3 x 3 x 3 mm^3). In each run, 305 functional volumes were acquired. For anatomical reference in each subject, a 192-slice T1-weighted 3D Magnetization Prepared Rapid Gradient Echo (MPRAGE) structural image was acquired in the same orientation as the Echo Planar Imaging (EPI) sequence (TR = 2.3 s, TE = 3.03 ms, flip angle = 9^0, FOV = 256 x 256 mm^2, slice thickness = 1 mm, acquisition matrix = 256 x 256, voxel size = 1 x 1 x 1 mm^3).

Stimuli and tasks

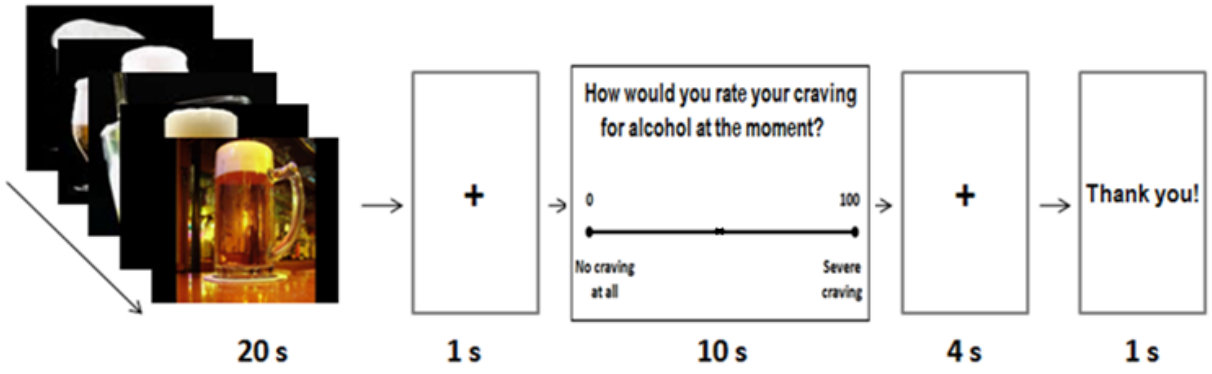


Figure 2.1. **Cue reactivity paradigm**.

An established cue-reactivity paradigm (Vollstädt-Klein et al., 2010) was conducted. In the block-designed fMRI task, 60 standardized alcohol-related pictures including 20 beer, 20 schnapps and 20 wine pictures, and 45 neutral pictures derived from the International Affective Picture System (IAPS) (Lang et al., 1999) were presented in a total of 20 pseudo-randomized blocks including 12 blocks displaying alcohol-associated stimuli and 08 blocks presenting neutral stimuli. Each block consisted of 5 randomized pictures which were displayed for 4 seconds each, resulting in a total duration of 20 seconds for a block. After every block, participants were asked to rate their desire to drink, i.e. craving for alcohol, on a visual analogue rating scale. Ratings ranged from 0 ("no craving at all") to 100 ("severe craving") and were recorded by pressing a button within a maximal time frame of 10 seconds. Subsequently, a black fixation cross and 'Thank you!' (4+1 seconds) was shown before a new picture block was started (Fig. 2.1). The total task duration was 12 minutes (refer to Nationales Genomforschungsnetz (NGFN)-Plus project).

Data pre-processing

The first 4 volumes of each run were discarded to remove the initial T1 magnetic transients in the data. After that, the remaining data were pre-processed using SPM8 (Wellcome Department of Imaging Neuroscience, London, United Kingdom: http://www.fil.ion.ucl.ac.uk/spm). First, all remaining scans were corrected for the timing differences between each slice and realigned to remove residual motion effects. The anatomical scan was co-registered to a mean re-sliced image obtained from the realign analysis. Next, the images were spatially normalized to the same standard space (MNI space) with voxel size 3 x 3 x 3 mm^3. Finally, the data were smoothed in space using a 6-mm full-width-at-half-maximum Gaussian filter (FWHM) to optimize the signal-to-noise ratio in small subcortical structures of interest as well as to reduce differences between activation images of subjects (Etzel et al., 2009).

Methods

In this study, each brain region was considered individually. The investigation for each region was conducted in the following two steps: (1) Constructing and collecting response patterns of the region from fMRI data recorded for the subjects (feature construction); (2) Classifying these patterns.

Step 1: **Feature construction**

1.1 Constructing and collecting response patterns for individual ROIs

Since our target was to find a feasible approach to the application of diagnostic functional imaging in clinical practice, the classification of a disorder or condition of the disorder for a subject (subject classification) was only the final consequence of the imaging inference process. Therefore, the response patterns of the brain whenever cues are exposed are our main object of interest. In the study, each block (B_i) was viewed as an independent observation of the brain response to the given stimulus. The response feature of the brain for each block (B_i) was expressed through its representative vector (volume) created by averaging over all scans measured within it (Fig. 2.2). As a result, for 12 blocks with alcohol-associated cues, each subject comprised 12 feature vectors, also considered as the response patterns of the brain to alcohol cues. Then, the response feature of a ROI k to alcohol cue for the block (B_i) was

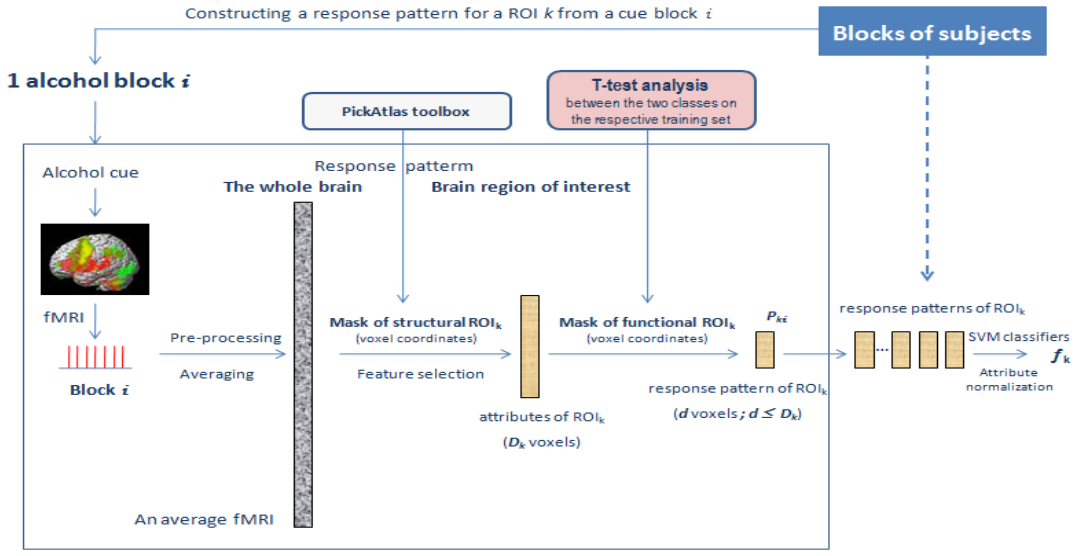


Figure 2.2. **Feature construction and collection of the response patterns for a ROI k.**

manifested by a feature vector (p_{ki}) extracted from the feature vector of the brain for that block using the mask of that ROI (Fig. 2.2). Similarly, from each subject for each ROI we collected 12 feature vectors, also considered as the response patterns of that ROI to alcohol cues. The feature vectors of the ROIs or the brain were interpreted as independent observations of the response patterns (also termed activation patterns) of the ROIs or the brain to alcohol cues and were used as input data for classifiers.

1.2 Normalizing the feature attributes of the response patterns

In the pre-processing step described earlier, all fMRI images measured for all the subjects were normalized spatially to the same standard space in order to minimize morphological variability between different subjects (normalization of voxel coordinate). In this step, before providing for classifiers as input, the data were normalized in the aspect of feature attribute (BOLD signal) to reduce the effect of large signal changes dominating those of smaller signal amplitude (normalization of voxel attribute) (Pereira et al., 2009). In this study, the method of scaling normalization was applied. For this, all input feature vectors ($\{p_{ki}\}$) of the training set for each classifier (f_k) of each ROI k were arranged in rows and columns (Fig. 2.2) in which each column was an input vector (each pattern) ($p_{ki} = \{a_{ki}^t\};\ t = 1, \dots, v$; v: size of the ROI k; i: the block where the pattern p_{ki} was measured; Fig. 4.3), and each row was an attribute of the vector ($\{a_{ki}^t\};\ i = 1, \dots, I_{training};\ I_{training}$: number of blocks (corresponding to the number of patterns) reserved for the training). The parameters of min and max value of each attribute (a_{ki}^t) were calculated only on the training set. Then, these parameters were used to scale all the attributes of

the vectors (p_{ki}) in both the training set and the testing set to the range [-1 +1] (($scaled\_a_{kj}^t = 2 * [\frac{(a_{kj}^t - min(\{a_{ki}^t\}))}{(max(\{a_{ki}^t\}) - min(\{a_{ki}^t\}))}] - 1$); $p_{kj} = \{a_{kj}^t\} \rightarrow scaled\_p_{kj} = \{scaled\_a_{kj}^t\}$; $i = 1, \ldots, I_{training}$; $j = 1, \ldots, I_{training+testing}$; $I_{training+testing}$: the number of the patterns for both training and testing). This scaling normalization was applied to normalize each attribute (row) (Fig. 2.2). In the study, the blocks with neutral stimuli were not considered.

1.3 Creating the mask of individual ROI

A ROI in fMRI analysis can be defined in terms of structural or functional properties. Structural ROIs are defined on the basis of anatomical structures, i.e. gyri, sulci, while functional ROIs are defined on the basis of data analysis obtained from experiments. Hence, for investigating brain activation, functional ROIs are our main objects of interest to ensure that the ROI only contains voxels that are truly activated under given stimulation (Etzel et al., 2009).

To focus observation on relevant brain regions with the aim that observation results are interpretable in clinical practice, for each brain region, its functional ROIs in this study were formed within its corresponding anatomical structures. The brain regions were chosen for the investigation based on the neurocircuitry of addiction proposed by Koob and Volkow (2010) including the VTA, VS, DS, thalamus, pallidum, amygdala, hippocampus, insula and PFC including ACC, orbital frontal cortex (OFC) and medial prefrontal cortex (mPFC) (Fig. 1.3).

1.3.1 Masks of structural ROIs

Structural ROIs were built with the Wake Forest University (WFU) PickAtlas toolbox version 2.4: http://fmri.wfubmc.edu/cms/software (Maldjian et al., 2003; Tzourio-Mazoyer et al., 2002) and then normalized to a standard space as same as the standard space of the smoothed data to create their corresponding structural ROI masks. The masks for the ACC, pallidum, thalamus, amygdala, hippocampus and insula were available in the toolbox. The others were built as follows: The

Table 2.1. **Size of structural ROIs with voxel size 3 x 3 x 3 mm^3**

ROI	Number of voxels
VTA	44
VS-DS	176
VS	88
DS	88
Pallidum	161
Thalamus	570
PFC	3136
ACC	987
mPFC	1499
OFC	1496
Amygdala	132
Insula	1215
Hippocampus	563
Whole brain	60588

PFC, OFC and mPFC were defined by Brodmann areas: PFC (8, 9, 10, 11, 44, 45, 46, 47); OFC (10, 11, 47) and mPFC (9, 10) (David & Mark, 2009; Goldstein & Volkow, 2011; Lee et al., 2006). The VS-DS, VS, DS and VTA were constructed based on small ROIs, each of which was defined by a sphere with 6 mm radius centered on the following MNI coordinates: right VS, [12, 15, −6]; left VS, [−12, 15, −6]; right DS, [12, 15, 6]; left DS, [−12, 15, 6] (Schacht et al., 2011); and VTA [0, -20, -12] (Eva et al., 2010) (Table 2.1).

1.3.2 Masks of functional ROIs formed within their corresponding structural ROIs

For each brain region, its functional ROIs were formed within its corresponding structural ROI (Fig. 2.2). For this, the scoring/filtering method was applied. The voxels of the structural ROI were ranked according to a given criterion (scoring), and then the top-ranked ones were chosen to form its corresponding functional ROI mask (filtering).

a. *Scoring with mass-univariate approach*

For a preliminary study, a mass-univariate approach was selected to score each individual voxel (a univariate test for a voxel). For this, a two-sided t-test analysis was conducted on a training dataset to specify the statistically different activation level of each voxel between the two groups of alcoholic patients and healthy controls characterized by a t-value. The underlying hypothesis for the t-test analysis is that there may be a different response to alcohol-associated cues and neutral cues in alcoholic patients whereas such a response may not occur in healthy subjects (Braus et al., 2001; Wrase et al., 2007; Park et al., 2007; Beck et al., 2009; Heinz et al., 2009). The analysis was done based on a general linear model which was implemented with SPM software version 8 (http://www.fil.ion.ucl.ac.uk/spm/software/spm8/).

First, a first-level analysis was performed within each subject (within-subject). In this study, the beer, schnapps and wine pictures used in the experiment counted as alcohol-associated cues indiscriminatingly. Each cue block was modelled as a boxcar function convolved with a canonical hemodynamic response function that began at the onset of the first cue of the block and ended at the end of the last cue. A high-pass filter (1/128 Hz) was applied to remove slow signal drift. A contrast image measuring the response difference between alcohol-associated stimuli and neutral stimuli was generated from the general linear model**.** Next, a two-sample ***t***-test at the second level between the two groups was conducted using random effect analysis

(between-subject). Here, each contrast image of the first-level was considered as a representative sample for its group. From this, an activation map (statistical map) for the whole brain was created, which described the different activation level of voxels between the two groups through a t-distribution; and an activation map for each structural ROI was also formed based on the activation map for the whole brain using its structural ROI mask.

b. *Filtering*

After ranking, one might still have to decide on how many of the top-ranked voxels should be chosen to form a functional ROI because the size of a functional ROI, i.e. the number of selected voxels, can have an impact on the classification performance (O'Toole et al., 2007; Etzel et al., 2009). One commonly used method is to threshold the statistical map and then to select the surviving voxels on the threshold within the observed brain structure (structural ROI). For a structure with size of *D* voxels, we observe simultaneously on *D* hypothesis testing results of the statistical map (one for each voxel). It is thus necessary to correct the threshold for multiple comparisons. For this, Bonferroni correction or false discovery rate control can be applied to define the threshold in order to ensure that the functional ROI only contains the voxels with their features deemed significant at a given level (Pereira et al., 2009). However, this approach can be quite sensitive to the specific threshold (Etzel et al., 2009). Especially in small structural ROIs, the approach may be problematic if no or few voxels are surviving.

For this reason and owing to the size difference of the investigated structural ROIs, in this preliminary study, we focused on a simple method in which the size of functional ROI was fixed at three different levels: 200, 100 and 50 voxels. A functional ROI with a given size *d* was defined as the *d*-voxels with the largest t-values identified within the corresponding structural ROI ($d \leq D$) using its activation map after ranked. In case the selected size to form the functional ROI was greater than the actual size of the corresponding structural ROI ($d \geq D$), the entire structural ROI was included ($d = D$), as was for instance the case for the VTA, VS-DS, VS, DS, amygdala and pallidum at a functional ROI size of 200 voxels (Table 2.1).

c. Sample for the t-test analysis

A classification task usually involves separating data into training and testing datasets (Figs. 2.5 & 2.6). Each pattern in the training set covers attributes of the pattern (p_{ki}) (i.e. voxel attributes)

and a class label (y_{ki}) assigned to it (p_{ki}, y_{ki}). Based on the training dataset, a given classification algorithm yields a corresponding model (classifier) to predict the class labels for the unlabeled patterns of the testing dataset only based on the attributes of the patterns (voxel attributes). Hence, to avoid overestimates of classification performance derived from peeking information of the testing data (Pereira et al., 2009), for each different loop of the training and testing, a separate t-test analysis at the second-level was conducted only on its corresponding training dataset of the loop. The creation of these loops will be presented in the next section of the evaluation.

Step 2: **Classifying the response patterns of individual ROI**

For each ROI k a separate SVM classifier f_k was used as an instrument to classify its response patterns (p_{ki}) into either class 1 (alcoholic class; $y_{ki} = f_k(p_{ki}) = +1$) or class 2 (control class; $y_{ki} = f_k(p_{ki}) = -1$) ($y_{ki} \in \{-1, +1\}$; p_{ki}: the response pattern of the ROI k recorded for block i). The training and testing for the SVM classifier was conducted as follow.

Training to obtain a model (SVM classifier) from labelled data

For a given labelled training dataset (p_{ki}, y_{ki}), there may exist many hyperplanes that separate the input vectors of this dataset (p_{ki}) into the two classes (Fig. 2.3a). Among the hyperplanes, there exists the optimal margin hyperplane with the largest margin of separation between the two classes. The vectors closest to the optimal margin hyperplane are called support vectors, and the distance between them and the hyperplane is called the margin of SVM classifier (Fig. 2.3b). In the cases where the input vectors are not linearly separable, they can be mapped into a (usually higher) dimensional feature

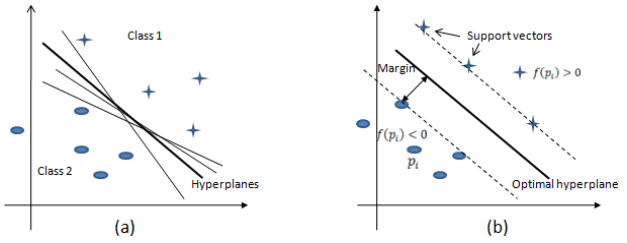


Figure 2.3. **Illustration for SVM classification.**
For linear SVM: $f_k(p_{ki}) = sign(w^T p_{ki} + b)$; For nonlinear SVM: $f_k(p_{ki}) = sign(w^T \emptyset(p_{ki}) + b)$; $f_k(p_{ki})$: decision function; w^T: weight vector; $\emptyset$: function mapping p_{ki} into a higher dimensional space; b: offset; p_{ki}: response pattern; *crosses*: the patterns of class 1; *circles*: the patterns of class 2.

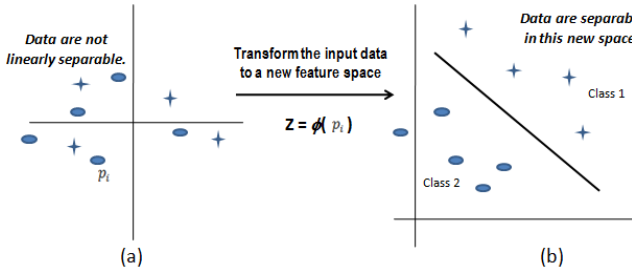


Figure 2.4. **Illustration for mapping data into a feature space.**

space ($p_{ki} \rightarrow \emptyset(p_{ki})$) in order that they become linearly separable in the projected space (Fig. 2.4). However, computation in the feature space can be costly because it is very high dimensional (typically infinite-dimensional). To solve this, the so-called kernel trick ($K(p_{ki}, p_{kj}) = \emptyset(p_{ki})^T \emptyset(p_{kj})$) is applied to avoid carrying out the mapping explicitly. Additionally, to improve to handle the data that are not fully linearly separable, one allows "errors" (ξ_i; $\xi_i \geq 0$) in classification by relaxing slightly the constraints to identify the optimal margin hyperplane i.e. $y_i(w^T \emptyset(p_{ki}) + b) \geq 1 - \xi_i$ instead of $y_i(w^T \emptyset(p_{ki}) + b) \geq 1$. The "errors" ξ_i are known as "slack variables" in optimization and provide an upper bound on the number of training errors. A SVM classifier using such a method is known as the "soft margin" SVM classifier; and a classifier that would generalize well on unseen data is then found by controlling both the classifier flexibility (Vapnik-Chervonenkis dimension via w) and the sum of the slacks ($\sum_{i=1}^{I_{training}} \xi_i$; $I_{training}$: the number of patterns for training) with the target of minimizing both of these two quantities. In the soft margin SVM, data points on the incorrect side of the margin boundary have a penalty that increases with the distance from it. To harmonize this issue, a regularization parameter (C) is used to control the trade-off between margin maximization (maximize $\frac{2}{\|w\|^2}$ *or* minimize $\frac{1}{2}\|w\|^2$) and training error minimization (minimize $\sum_{i=1}^{I_{training}} \xi_i$). Specifically, the SVMs require the solution for the following optimization problem (a):

$$\underset{w,b,\xi}{\text{minimize}} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{I} \xi_i$$

$$\text{subject to} \quad y_i(w^T \emptyset(p_{ki}) + b) \geq 1 - \xi_i$$

$$\forall\, \xi_i \geq 0;\ i = 1 \rightarrow I_{training}\text{: the number of patterns for training}$$

When designing a SVM classifier, the first task is to select kernel ($K(p_{ki}, p_{kj})$) and regularization parameter (C), and for a given kernel, to set the parameters that the kernel function may depend on e.g. width of a Gaussian kernel (corresponding to Radial Basis Function (RBF)) (γ; $K(p_{ki}, p_{kj}) = e^{-\gamma\|p_{ki} - p_{kj}\|^2}, \gamma > 0$) or degree of a polynomial kernel (d; $K(p_{ki}, p_{kj}) = (\gamma p_{ki}^T p_{kj} + r)^d$). These parameters are called hyper-parameters and have considerable effects on the classifier flexibility and training error. Based on training dataset together with the defined kernel/hyper-parameters, the values of parameters for specifying the optimal margin hyperplane

i.e. b and w or α_i (Lagrange multipliers for dual formulation (substituting dual optimization problem for the objective problem (a)); $w = \sum_{i=1}^{I_{training}} y_i \alpha_i \emptyset(p_{ki})$; $\alpha_i > 0$: points on the margin (support vectors)) are estimated to produce a model for classification (SVM classifier). Since different datasets can be compatible with different kernels (Hsu C-W, Chang C-C & Lin C-J, 2010), one might try all kinds of kernels/hyper-parameters, and choose the kernel with the best performance. Nonetheless, this may cause overfitting (Ben-Hur & Weston, 2010).

In this study, 14 different brain structures (structural ROI; Table 2.1) with three different sizes of the functional ROIs (200, 100 and 50 voxels) corresponding to each structure were observed. Thus, significant differences among the datasets of the response patterns collected for the observations are likely. This implies that for some observations the datasets are compatible with a given kernel, however for the others, there may be no compatibility. In order to take advantage of the compatibility of each kernel (e.g. either linear or nonlinear kernel) with different training datasets, both of the linear and radial basis function kernel were applied interchangeably. The selection of the kernel/hyper-parameters to form a particular classifier f_k from a given training set was done as follows. First, for each ROI k, the values of the hyper-parameters (C for the linear kernel and C, γ for the RBF kernel) for the classifier f_k were specified via grid search using "grid.py" module with 5-fold cross-validation and the exponentially growing sequences of grid parameters: $C = 2^{-5}, 2^{-3}, \ldots, 2^{15}$; $\gamma = 2^{-15}, 2^{-13}, \ldots, 2^{3}$ (Hsu C-W, Chang C-C & Lin C-J, 2010). After that, for each loop of training and testing (L^t), the classifiers with different pairs of the specified kernels/hyper-parameters were trained and tested in turn on the 45 nested loops created within its corresponding training dataset ($nested\_L_j^t; j=1 \rightarrow 45$; Fig. 2.5 & 2.6; see the next section on creating examples). Based on the result of this testing, the classifier that has yielded the highest average accuracy was selected for the loop (L^t). The Matlab implementation of the LibSVM version 3.1 (http://www.csie.ntu.edu.tw/~cjlin/libsvm/) was used in this study. For a more detailed description of the SVM formalism, refer to the machine learning literature (Boser et al., 1992; Vapnik 1998; Schölkopf & Smola, 2000; Vapnik, 2000; Wang, 2005).

Testing a SVM classifier (model) on unseen data

For testing on an unlabeled pattern ($\dot{p}_{ki}$) of testing dataset, the distance of the pattern ($\dot{p}_{ki}$) from the hyperplane was calculated using the classifier f_k which has been specified from the training.

The decision to classify the pattern ($\dot{p}_{ki}$) into either class 1 or class 2 was determined by its sign ($y_{ki} = sign(w^T \phi(\dot{p}_{ki}) + b)$). This means that if $f_k(\dot{p}_{ki}) > 0$, the pattern is classified into class 1 ($y_{ki} = +1$), and conversely, if $f_k(\dot{p}_{ki}) < 0$, it is classified into class 2 ($y_{ki} = -1$).

Evaluation

Creating examples

Cross-validation (CV) is a statistical method used to evaluate or compare learning algorithms by repeatedly dividing data into two different datasets: one used to train a model and the other to measure its classification performance (Payam et al., 2009). Here, the 107 subjects who were included in the study were randomly divided into an 80-subject cross-validation (CV) sample

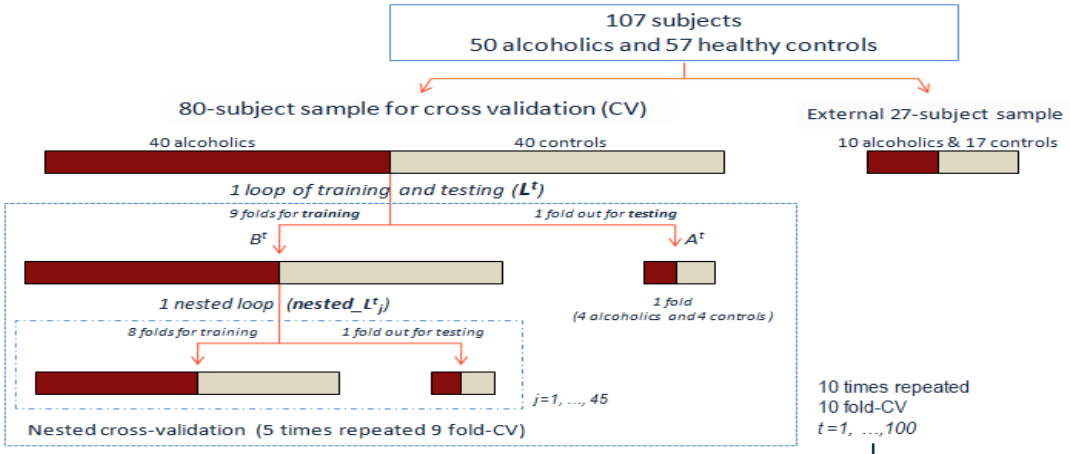
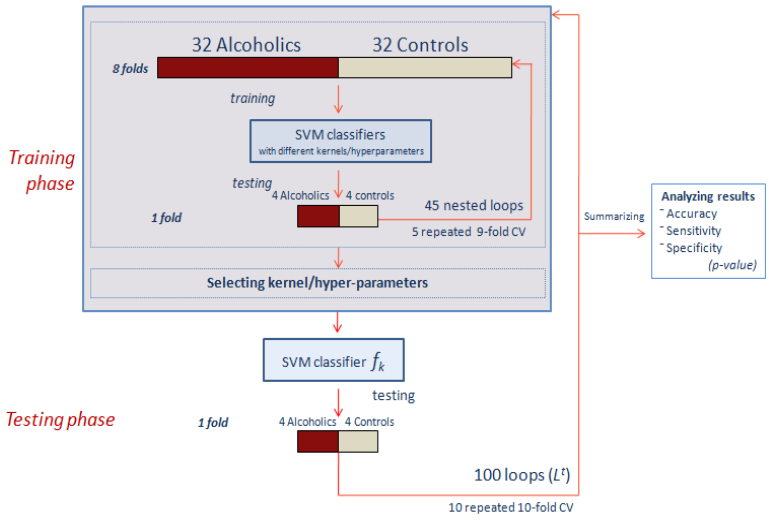


Figure 2.5. **Creating examples for the evaluation**

(including 40 alcoholics and 40 controls) and a 27-subject sample (10 alcoholics and 17 controls) for external validation.

On the 80-subject sample, a stratified 10-fold cross-validation procedure repeated 10 times was applied to create 10 rounds for the evaluation. Firstly, the 80-subject sample was



Loops of training and testing for the evaluation on the 80-subject sample

partitioned randomly into 10 equal folds, each of which included 8 subjects (4 alcoholics and 4 controls, corresponding to 96 response patterns of the brain or ROI collected from them). Based on the 10 folds, each round created 10 loops of training and testing such that within each loop (L^t; t =1→10) a different fold of the data containing 8 subjects (A^t) was held-out for testing while the remaining 9 folds including the 72 other subjects (B^t) were used for training. To eliminate sampling bias, this procedure was repeated 10 times on random cross-validation splits of the 80-subject sample (L^t; $t = 1 \rightarrow 10$ x 10). Moreover, within the training segment (B^t) of each loop (L^t), a 9-fold cross-validation procedure repeated 5 times to create 45 nested loops of training and testing ($nested\_L_j^t$; $j = 1 \rightarrow 5$ x 9) was applied for adjusting the kernels/hyper-parameters of the SVM classifiers described in the previous section (Fig. 2.5).

Evaluating classifiers

As mentioned earlier, for each brain region with a defined size of functional ROI, a corresponding classifier would be formed. We used the 80-subject sample with the cross-validation procedure to evaluate their classification performance in terms of accuracy, sensitivity and specificity. Accuracy, sensitivity and specificity were averaged over the testing results of the 100 test datasets created from the 10-fold cross-validation repeated 10 times. Each test comprised the 96 patterns collected from the 8 subjects of the corresponding testing dataset. Accuracy was defined as the relative number of total samples (or patterns) classified correctly. Sensitivity was defined as the percentage of correctly classified patterns into alcoholic class. Specificity was defined as the percentage of correctly classified patterns into control class.

Result significance for a classifier

The statistical significance of the classification result for a classifier was analyzed based on rejecting the null hypothesis. The null hypothesis assumes that there is no difference between the samples of the two classes. The p-value to reject this hypothesis was estimated by cumulative probability function (p-value = $P(X \geq k)$) where X is a random variable with a binomial distribution with N trials corresponding to the number of classified samples, k successful trials and the probability of a successful trial p = 0.5 for the two classes. Each test sample (each pattern) was regarded as an independent Bernoulli trial (Pereira et al., 2008). To avoid the optimistic evaluation in multiple testing, the p-value was adjusted using the Bonferroni

correction with 1% significant level (Bland & Altman, 1995; Pereira et al., 2008). Under the Bonferroni correction, the accuracies of classifiers were said to be significant with 1% level (p_{corr}<0.01) if their uncorrected p-values were smaller than the corrected alpha (α=0.01/(N$\times m$); N: number of trials (samples or classified patterns); *m*: number of observations).

Comparison of two different classifiers

The comparison of two different classifiers on the same dataset e.g. f_k and f_l in terms of statistical significance was carried out using McNemar's test (Dietterich, 1998; Roggo et al., 2003; Jaber et al., 2010). This test was based on the chi-square distribution (χ^2) with one degree of freedom. Accordingly, the critical value with 5 % significance level is 3.8415 ($\chi^2_{(1,0.95)}$). The null hypothesis assumes that the two classifiers f_k and f_l would have no difference of error rate. Then the null hypothesis is rejected if McNemar's value (χ^2_M) is greater than 3.8415, and these two classifiers are said to be significantly different (p<0.05). In this study, McNemar's value was calculated using the following formula:

$$\text{McNemar's value} = \frac{(|\mathrm{n}_{01} - \mathrm{n}_{10}| - 1)^2}{\mathrm{n}_{01} + \mathrm{n}_{10}}$$

Where n_{01} was the number of samples incorrectly classified by f_k but correctly classified by f_l, and n_{10} was the number of samples correctly classified by f_k but incorrectly classified by f_l.

To get a more robust analysis, the prediction models (or classifiers) which have been tested on the 80-subject sample were again evaluated on the external dataset including the 27 subjects. Since this dataset was up to then never used in the previous design process, the results on this dataset were not subject to any model selection bias.

Evaluating correlation between the t-value and performance of classifier

In order to assess validity of using the t-values at the second-level statistical analysis for the formation of functional ROIs within their corresponding structural ROIs, we investigated effects of the t-values on the classification performance of classifiers or in other words, whether the performance of classifiers for the functional ROIs with the high t-values is better than those with the lower t-values. For this, a second level t-test analysis that shows difference of activation between alcoholics and controls on the 80-subject sample was conducted to create a map of the t-

value distribution for the whole brain and correspondingly, a map of the second-level t-value distribution for each structural ROI using its structural mask. As described earlier, a map of the t-value distribution for each functional ROI was specified based on a ranked map of the t-value distribution of the corresponding structural ROI and given size *d*. Then, the comparison of t-value distribution between the two different functional ROIs was analyzed using a t-test analysis for difference between two means of these two functional ROIs in the case of independent samples (Zikmund et al., 2013). In combination with this comparison, the comparison of classification performance between their two corresponding classifiers was also analyzed to provide answers for the assessment.

Results

1. Performance on the 80-subject cross-validation sample

The results were summarized from testing on 100 sample sets that were created from 10 repeated 10-fold cross-validations (8 subjects for each test; 12 response patterns of the brain or ROIs collected from 12 cue blocks for each subject). As a result, the total of the classified patterns for each observation on each functional ROI was 9600 (N = 8 x 12 x 100). For evaluation of the significance of classifiers in the observations, their accuracies were said to be significance with 1% level under the Bonferroni correction (p_{corr}<0.01) if their uncorrected p-values were smaller than 0.248×10^{-7} (0.01/(N$\times m$); N: number of trials (classified patterns); N = 9600; *m*: number of observed classifiers in the same context; *m* = 42; Tables 2.2, 2.3 and 2.4). According to the cumulative probability function on a binomial distribution, this also means that the accuracy of a classifier is significant with p_{corr}<0.01 if it is greater than 52.8% (uncorrected p-value = 0.248×10^{-7}; N = 9600; p = 0.5; p: probability of success in one trial).

a. *Functional ROI with the size of 200 voxels*

In the second-level analysis, the functional ROIs such as the PFC were formed by the voxels with more significantly different activation (t-value = 1.79 ± 0.35) than those of ACC (t-value = 1.45 ± 0.4) and VS (t-value = 0.3 ± 0.54) (z = 5.55 and 22.34; *p*<0.01). In contrast to this, the ACC and VS yielded the classification outperformance (62.8% and 60.3% vs. 51.8% accuracy respectively; $\chi_M^2 =$ 23.61 and 9.37; *p*<0.01; Table 2.2).

Table 2.2 **The performance of pattern classification for functional ROIs with the size of 200 voxels**

ROI N=9600	t-value distribution at the 2nd level analysis				Classification performance (%)		
	max	*min*	*mean*	*sd*	*Accuracy*	*Sensitivity*	*Specificity*
VTA	1.69	0	0.45	0.5	55.7	57.9	53.6
VS-DS	2.39	0	0.51	0.58	58.8	61.1	56.4
VS	1.77	0	0.3	0.54	**60.3**	61.4	59.2
DS	2.39	0	0.73	0.53	50.7	46.9	54.6
Pallidum	2.37	0	0.35	0.6	59.8	61.9	57.7
Thalamus	2.58	0.74	1.11	0.31	56.7	55.2	58.2
PFC	3.12	1.4	1.79	0.35	**51.8**	54.3	49.3
ACC	2.45	0.87	1.45	0.4	**62.8**	64.9	60.7
mPFC	3.01	1.08	1.54	0.36	45.9	43.3	48.5
OFC	3.12	0.37	1.01	0.55	45.0	45.5	44.6
Amygdala	0.21	0	0.05	0.04	53.7	58.0	49.4
Insula	2.98	1.19	1.66	0.39	48.2	50.8	45.5
Hippocampus	2.3	0.27	0.98	0.5	51.9	54.7	49.1
Whole brain	4.21	2.56	2.83	0.26	61.8	66.3	57.3

b. Functional ROI with the size of 100 and 50 voxels

The reduction of defined size for the functional ROIs from 200 voxels to 100 and 50 voxels increased the distribution of the higher t-values correspondingly e.g. for the ACC with the size of 200, 100 and 50: t-value = 1.45 ± 0.4, 1.76 ± 0.3 and 2.02 ± 0.19 respectively; z = 4.18 and 6.02; $p<0.01$. However, the increases did not yield a correspondingly better performance (62.8%, 60.6% and 59.5% accuracy respectively; Tables 2.2 & 2.3).

Table 2.3 **The performance of pattern classification for functional ROIs with the size of 100 and 50 voxels**

ROI N=9600	Functional ROI size of 100 voxels					Functional ROI size of 50 voxels				
	t-value distribution		Classification performance (%)			t-value distribution		Classification performance (%)		
	mean	*sd*	*Accuracy*	*Sensitivity*	*Specificity*	*mean*	*sd*	*Accuracy*	*Sensitivity*	*Specificity*
VTA	0.45	0.5	55.7	57.9	53.6	0.45	0.5	55.7	57.9	53.6
VS-DS	0.9	0.5	57.0	55.7	58.4	1.3	0.33	59.2	56.6	61.8
VS	0.3	0.54	60.3	61.4	59.2	0.53	0.64	63.9	65.0	62.9
DS	0.73	0.53	50.7	46.9	54.6	1.1	0.39	54.9	48.9	60.9
Pallidum	0.56	0.67	63.7	65.3	62.1	1.11	0.55	60.8	63.4	58.1

Thalamus	1.35	0.27	52.6	53.8	51.4	1.55	0.25	50.4	52.2	48.6
PFC	2.05	0.32	47.8	50.6	45.1	2.28	0.29	50.5	50.4	50.6
ACC	1.76	0.3	**60.6**	61.3	60.0	2.02	0.19	**59.5**	58.9	60.1
mPFC	1.81	0.32	47.9	47.7	48.1	2.05	0.29	49.5	48.4	50.7
OFC	1.41	0.5	47.0	49.8	44.1	1.78	0.47	48.2	46.9	49.4
Amygdala	0.06	0.03	58.0	61.7	54.4	0.08	0.03	56.7	61.5	51.9
Insula	1.96	0.32	49.5	52.5	46.6	2.2	0.29	53.6	56.9	50.2
Hippocampus	1.4	0.36	45.8	48.2	43.5	1.7	0.24	44.2	45.6	42.8
Whole brain	3.17	0.28	58.6	61.2	56.0	3.17	0.28	55.2	57.5	52.9

2. Performance on the external 27-subject sample

For testing the classifiers for the ROIs on the external 27-subject sample, we trained the classifiers on the 80-subject cross-validation sample and then tested them on the external sample, and only focused on the functional ROI size of 200 voxels. The results showed that the VS, ACC still kept the significantly high performance on the external 27-subject sample (62.3% and 67.6% accuracy respectively; p_{corr}<0.0001; N = 27 x 12; m = 14).

Table 2.4 **The performance of pattern classification on the external 27-subject sample**

ROI N = 324	**Classification performance (%) For functional ROI size of 200 voxels**		
	Accuracy	*Sensitivity*	*Specificity*
VTA	52.5	48.3	54.9
VS-DS	58.3	68.3	52.5
VS	**62.3**	59.2	64.2
DS	72.8	63.3	78.4
Pallidum	54.9	49.2	58.3
Thalamus	60.5	24.2	81.9
PFC	57.4	65.8	52.5
ACC	**67.6**	51.7	77.0
mPFC	71.6	80.0	66.7
OFC	67.6	71.7	65.2
Amygdala	34.9	20.8	43.1
Insula	65.7	52.5	73.5
Hippocampus	29.6	33.3	27.5
Whole brain	58.6	44.3	67

Discussion

Mass-univariate approach for the formation of a functional ROI from its structural ROI

The investigation on individual ROIs indicated that the brain regions of ACC and VS play a prominent role in the diagnostic classification of alcohol dependence using fMRI, which can be seen by the good classification performance achieved on the regions (ACC, 62.8% accuracy (Table 2.2); VS, 63.9% accuracy (Table 2.3); p_{corr}<0.0001). This result is compatible with previous studies on cue reactivity in alcohol dependence, in which the VS and ACC has exposed significant activation under impact of alcohol-associated stimuli in alcoholic patients compared to those of controls (Braus et al., 2001; Heinz et al., 2004, 2009; Wrase et al., 2007; Park et al., 2007; Beck et al. 2009; Schacht et al., 2011). However, the investigation on the second-level t-value distribution of these regions with the functional ROI size of 200 voxels (ACC, t-value = 1.45 ± 0.4 and VS, t-value = 0.3 ± 0.54) showed no corresponding prominence compared to those with the significantly lower performance of classification such as the PFC (t-value = 1.79 ± 0.35) (62.8% and 60.3% vs. 51.8% accuracy respectively; $\chi^2_M = 281.42, 154.37$; p<0.0001). The differences of t-value distribution between the functional ROIs e.g. between the VS and PFC (z = 22.34; p<0.00001) can be derived from the significant differences of analyzed structural ROI size e.g. the VS, 88 voxels and the PFC, 3136 voxels (Table 2.1). Nonetheless, the evidence indicates that the functional ROIs formed from a set of voxels with the better t-values may not yield a correspondingly better classification performance using SVM classifier. The inference is in line with the previous results of fMRI classification showing that the classification performance with feature selection using the univariate approach was lower than that using a multivariate approach such as SVM, Gaussian Naive Bayes or Linear Discriminant Analysis (Pereira et al., 2009). These results therefore appear to confirm that using the t-value as an indicator to rank voxels for the formation of functional ROIs within their corresponding structural ROIs is not an optimal method of feature selection for the classification using multivariate methods such as SVM.

How to form a functional ROI from its corresponding structural ROI appropriately?

To find clues in answer to this question, let's reconsider the way that a SVM classifier makes a decision and the nature of classified data.

Decision-making of a SVM classifier

With a multivariate approach, the decision-making of a SVM classifier for a classified pattern is influenced by individual variables (e.g. voxel attributes of the pattern (p_{ki}) and their corresponding weights (w)) and correlation between them characterized by a mathematical formula connecting the variables together (e.g. $f(p_{ki}) = sign(w^T p_{ki} + b)$; Fig. 2.3). This means that the SVM classifier is appropriate for recognizing informative patterns rather than contrast differences in individual voxels (Norman et al., 2006; Pereira & Botvinick, 2011). Apart from the evidence of loose correlation between the performance and t-value for the investigated ROIs discussed above, the argument is supported by the result of a study by Mourão-Miranda et al. (2005) indicating that there was a difference in the selected voxel set over a defined threshold between a univariate and a multivariate approach when analyzing difference of activation between tasks. This suggests that the formation of a functional ROI to gain a high classification performance using a SVM classifier should be based not only on the characteristics of each individual voxel (voxel attribute) but also on those of correlation between the voxels within the observed structure.

Nature of classified data

For this reason, in order to answer the question above, let's try to reconsider the nature of the information that we are working on. As mentioned in the literature review, the nature of our classified data is vascular changes represented by BOLD signals and transformed into voxel attributes (image resolution value; Fig. 1.6). The vascular changes indirectly reflect activity of the brain under impact of given stimulation (Mathews, 2001; Logothetis et al., 2001; Logothetis & Pfeuffer, 2004). The neurovascular coupling mechanisms are thought to relate to one or more vasoactive mediators such as nitric oxide, adenosine and changes in K^+ or hydrogen ions (i.e., pH) which are released from active nerve fibers to mediate local metabolism and blood flow (Yang et al., 2011). Furthermore, single neurons do not work independently but rather function in large aggregates (Mathews, 2001). Consequently, the BOLD signals or the data on which we are working appear to be operated by a mechanism involving the functional specialization of each brain region according to which the greater the homogeneity of structure and function, the stronger the correlation between the components within that region. This mechanism is evidenced by the result of the study of brain complexity measurement that indicated stronger and

more complex local connections between neurons of similar specificity forming neuronal groups than those between neurons of the different functional groups (Tononi, 1994). Hence, the correlation between voxels within an investigated brain region can be considered as degree of effective connectivity between these voxels in response to given stimulation which complies with this operating mechanism. The interpretation is supported by the classification results of PFC and ACC. The anatomical structure of PFC is formed from the substructures including the ACC, mPFC and OFC, so obviously for the same size of selected voxels, the functional ROIs of PFC possesses the sets of voxels with the larger t-values than those for its substructures such as the ACC (e.g. for the size of 200 voxels, the PFC, t-value = 1.79 ± 0.35 vs. the ACC, t-value = 1.45 ± 0.4; z = 5.55; p<0.0001; Table 2.3). Nevertheless, in contrast to this, the classification performance on the functional ROI of PFC was significantly lower than those on the functional ROI of ACC (51.8% vs. 62.8% accuracy respectively; $\chi_M^2 = 23.61$; p<0.0001; Table 2.2). Similar evidence was also found on the structures of VS-DS and VS, where the VS-DS was formed by combining the VS and DS. This combination did not yield a better classification performance for the VS-DS than that for the VS (58.8% vs. 60.3% accuracy respectively; $\chi_M^2 = 9.12$; p=0.0025; Table 2.2).

Further evidence to support this argument can be observed in the analysis of correlation between classification performance and ROI size. The results indicate that the reduction in size of functional ROI to filter out more voxels with no different activation can yield better performance for several brain regions e.g. the pallidum, 59.8% vs. 63.7% accuracy for the sizes of 200 and 100 voxels respectively. However, this was not recorded for the others e.g. the ACC, 62.8%, 60.6% and 59.5% accuracy for the sizes of 200, 100 and 50 voxels respectively (Tables 2.3 & 2.4). These results suggest that the size of a functional ROI may not be a critical factor influencing its performance of classification but rather the performance may essentially depend on its characteristics of structure and functional specialization as well as its role in the investigated disorder. The inference is evidenced by the classification results of the brain regions of interest in the case of analyzed functional ROI size with 200 voxels (Table 2.2) showing the significantly higher performance of classification for the core brain regions in the pathophysiologic mechanism of alcohol dependence e.g. the VS (88 voxels) than those for the others e.g. the amygdala (132 voxels) (60.3% vs. 53.7%; $\chi_M^2 = 94.74$; p<0.0001; Table 2.2).

The above interpretation is in accordance with the evidence showing the "searchlight" approach which is operated on correlation between neighbor voxels proved more sensitive than the mass-univariate approach in identifying brain regions containing category information between tasks (Kriegeskorte et al., 2006). Further, the outperformance for the VS and ACC compared with that for the VS-DS and PFC recorded on the external sample was consistent with that on the cross-validation test (the VS, 62.3% vs. VS-DS, 58.3% accuracy; $\chi^2_M = 2.72$; p=0.0992; and the ACC, 67.6% vs. PFC, 57.4% accuracy; $\chi^2_M = 6.36$; p=0.0117; Table 2.5). Since the external sample had never been previously used in the process of model building, this shows that these findings are not a result of model selection bias.

Observation on the whole brain

As just discussed, the formation of the PFC including the ACC, mPFC and OFC or the VS-DS combining the VS and DS did not yield the functional ROIs with the better classification performance than that of some of their corresponding substructures. Thus, for a broader interpretation where the whole brain is considered as a single structure, it is reasonable that the observation on the whole brain can be appropriately separated into multiple observations on relevant brain regions to get better benefits of performance and interpretation of classification result (e.g. the whole brain, 61.8% vs. the ACC, 62.8%; the VS, 63.9% accuracy; $\chi^2_M = 2.39, 11$; p=0.1219, 0.0009 respectively; Tables 2.2 & 2.3). In other words, it assists the practicality of deeper focus on the characteristics of structure and function of the investigated brain regions using multivariate methods such as SVM on fMRI data. Further, in a context where the combination of different brain structures into a single structure may not yield a better performance than that of some of its individual substructures, the questions are raised whether and how to combine multiple observation results on several relevant brain regions is valid for a fMRI classification. These issues will be elucidated in the next chapter.

CHAPTER III

fMRI CLASSIFICATION BASED ON MULTIPLE LINES OF EVIDENCE

Introduction

The result of the first study indicates the inefficiency of using the t-value to form the functional ROIs from their corresponding structural ROIs. Also, it suggests that the classification performance for the ROIs may depend on their role in the pathophysiological mechanism of investigated disorder much more than on the different activation level of individual voxels. From this basis, the question arises that if a functional ROI is just its corresponding anatomical structure, whether it can give an expected result. If this is validity, it can give a very important advantage in looking for a way to apply diagnostic functional imaging in practice because the voxel set of each functional ROI is then the same in all the tests and in all the subjects.

Additionally, for an imaging diagnosis in practice, a decision of image recognition is usually more confident if it is based on a synthesis of multiple observation results on multiple brain regions than if it is only focused on a single brain region. However, whether this approach is feasible in relapse prediction in alcohol dependence using fMRI remains a question. The clarification of the issues was the objective of our second study. Specifically, this study was to determine the validity of deeper focus on the anatomical structures of brain regions of interest in relapse prediction as well as the validity of prediction combining individual observation results on relevant brain regions using fMRI.

Materials and methods

Materials

Participants

The study was conducted on 40 patients (including 20 relapsers and 20 abstainers) recruited randomly from the 50 alcoholic patients who were included in the first study. Apart from those described in the section on the materials in the first study, before fMRI experiment, all the

patients were assessed with regard to the severity of their alcohol dependence using the Alcohol Dependent Score (ADS) (Skinner & Horn, 1984). Severity of alcohol craving was measured with the Alcohol Urge Questionnaire (AUQ) (Bohn et al., 1995) and the Obsessive Compulsive Drinking Scale (OCDS) (Anton et al., 1995). After discharge, the patients were interviewed about their alcohol consumption over a follow-up period of another six months, and were classified as abstainers (no alcohol consumption; n = 20) and relapsers (any alcohol consumption; n = 20).

Table 3.1. **Clinical data of the two groups of relapsers and abstainers**

Characteristics	Relapsers n = 20		Abstainers n = 20		*p value*
	mean	*sd*	*mean*	*sd*	
Age (y)	44.1	11.9	44.4	10.9	*0.76*
Age at onset (y)	36.4	8.8	36.2	9.6	*0.48*
Sex (male : female)	14:6		15:5		*0.49*
ADS	15.4	6.7	16.1	6.8	*0.53*
OCDS	16.5	5.6	15.2	5.7	*0.87*
AUQ score	21.5	7.8	26.9	22.8	*0.09*
Number of cigarettes per day	19	14.6	18.7	11.4	*0.20*
Number of abstinent days before fMRI	12.3	4.8	11.6	5.5	*0.88*
Number of abstinent days until relapse	61	47.0			
Alcohol intake during follow-up period (g)	6494	8204			

Data pre-processing

The pre-processing in the second study was similar to that in the first study.

Methods

The classification algorithm was designed in a way that emulated the way clinicians usually use to diagnose imaging in practice. First, to classify a subject, they observe and recognize the features of several individual brain regions of interest. Second, they infer the feature of the brain from these observation results. Finally, based on such multiple observations they make a diagnosis of disorder or condition of the disorder for the subject who has produced the images. In the study, for the first step, the observation and recognition of the response patterns of individual brain regions were done by SVM classifiers (*step A 1 including A 1.1 and A 1.2*; Fig. 3.1). For the second step, the inference based on these observation results was done using Bayesian

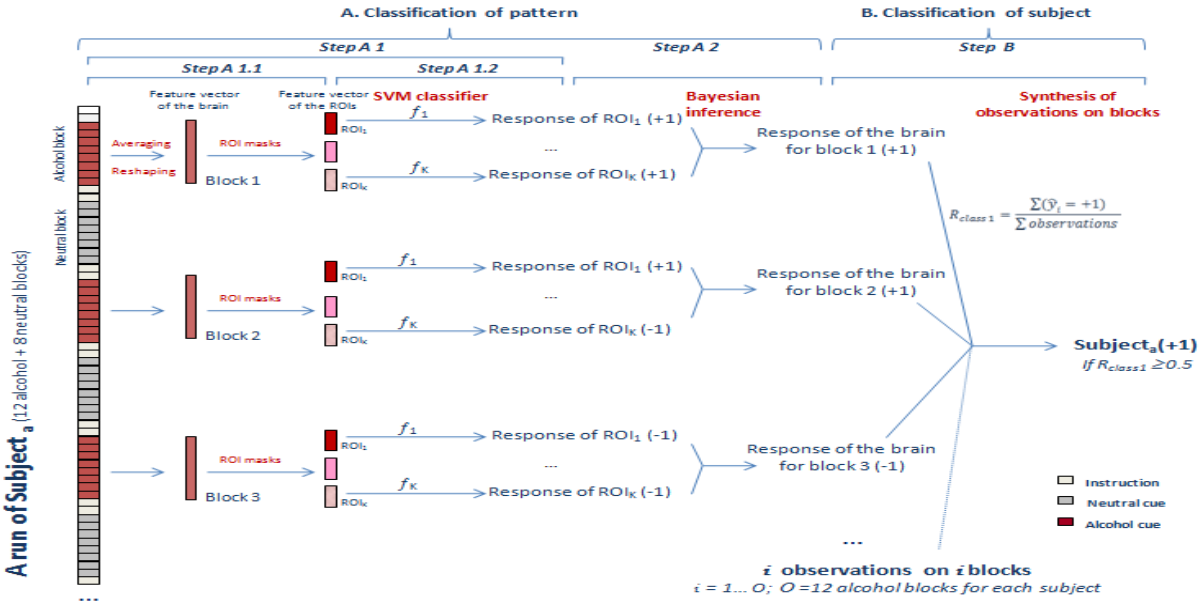


Figure 3.1. **Illustration of general classification algorithm for a particular subject$_a$** (+1: relapse class; -1: abstainer class).

inference (*step A 2*). The prior experience for the inference was learned from the training set. These two steps were seen as the step of pattern classification (*step A*). The final step, the step of subject classification, was considered as the synthesis step of various observation results of the response patterns of the brain whenever cues are exposed in order to make a prediction for the subject who has produced these patterns (*step B*).

A. Classification of pattern

A 1: Observation on individual ROIs

A 1.1: Feature construction

The method of constructing the response patterns of investigated brain regions was applied in a manner similar to that used in the first study. As a result, for a particular ROI k, each subject provided 12 response patterns (p_{ki}); and for 20 relapsers and 20 abstainers, 480 response patterns were collected and interpreted as 480 independent observations of response images of the ROI

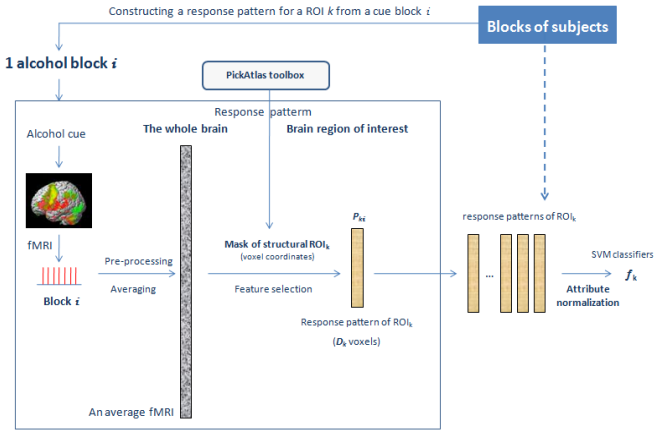


Figure 3.2. **Feature construction for a ROI *k* without the t-test analysis at the second level between the two classes.** A functional ROI was its corresponding structural ROI.

k to alcohol cues. They were used as input data of a classifier f_k. However, in this study there were the two following important differences compared to those in the first study.

(1). *Formation of functional ROI*:

The investigation was focused more deeply on anatomical structures of the ROIs. The anatomical structures of each brain region in the left, right and both hemispheres were investigated separately (Table 3.2). While in the first study, functional ROIs were formed within their own structural ROIs using the scoring/filtering method based on the t-value obtained from the statistical analysis at the second-level on the training dataset, in this study the functional ROIs were just their corresponding structural ROIs (Fig. 3.2).

Table 3.2. **Size of structural ROIs with voxel size 3 x 3 x 3 mm^3**

ROI	Number of voxels		
	Left	*Right*	*Bilateral*
VTA	12	16	44
VS	44	44	88
DS	44	44	88
Pallidum	79	79	161
Thalamus	282	288	570
ACC	498	489	987
mPFC	681	818	1499
OFC	704	792	1496
Amygdala	63	69	132
Insula	612	603	1215
Hippocampus	281	282	563

(2). *Normalization of the feature attributes of patterns*:

Besides the scaling normalization for the response patterns mentioned in the section on feature construction in the first study (method 1), we investigated the second method combining both of the scaling normalization and z-score normalization (method 2) to reduce variability of the patterns between blocks as well as subjects in situations where it cannot be done using the scaling normalization alone. For this method, before applying the scaling normalization, each input vector (p_{ki}) was additionally z-score normalized to have mean 0 and standard deviation 1. For each vector (p_{ki}), the average of all the attributes of this vector was subtracted from each of its attributes, and then the result of the subtraction was divided by the standard deviation of its attributes ($zscored\_a_{ki}^t = \frac{a_{ki}^t - mean(p_{ki})}{std(p_{ki})}$; $p_{ki} = \{a_{ki}^t\} \rightarrow zscored\_p_{ki} = \{zscored\_a_{ki}^t\}$; $t = 1, \dots, v$; v: size of the ROI k; Fig. 4.3). This z-score normalization was applied to normalize each individual pattern or each vector (column; Fig. 3.2).

A 1.2: *Classifying the response patterns of individual ROI*

A SVM classifier was used as an instrument to classify the response patterns of an individual ROI in the same manner as those in the first study. Hence, for each ROI k, a separate SVM

classifier f_k was used to classify its response patterns (p_k) into either class 1 (relapser class; $f_k(p_k) = +1$) or class 2 (abstainer class; $f_k(p_k) = -1$) (Fig. 3.1).

A 2: Combination of the observation results on multiple ROIs

Step *A 1* was seen as the step to observe and identify individually the response patterns of brain regions of interest to the given stimulation recorded within blocks. With the design described in the first study, for each block B_i a response pattern of the whole brain corresponding to the average volume over all the scans recorded for that block was constructed. From this pattern of the brain, a pattern of each ROI for block B_i was extracted using its ROI mask. Thus, the response pattern of the brain for each block can be considered as an overall picture which is put together by the pieces of the response patterns of brain regions extracted from it. The combination of the observation results on these pieces can help us infer the response pattern of the brain for that block (also termed 'block classification') (Fig. 3.1).

Instrument for the combination

Bayesian inference was used as an instrument for this combination. This inference derived from Bayes' rule has been applied to solve uncertainty. Bayes' rule shows the relation between two conditional probabilities that can infer each other (Hall, 2012). Considering a hypothesis H, its alternative hypothesis $\overline{H}$, and the observation E, the posterior probability of a hypothesis H after observing E is given by

$$P(H|E) = \frac{P(E|H) \cdot P(H)}{P(E)},$$

where $P(H)$ is the prior probability of H before observing E, $P(E|H)$ is the probability of a the observation given that the hypothesis H is true, and $P(E) = P(E|H) \cdot P(H) + P\left(E\middle|\overline{H}\right) \cdot P\left(\overline{H}\right)$.

In this inference, Bayes' rule is applied to update the probability estimate for a hypothesis after evidence has been observed (Tipping, 2010). Specifically, in this study, the inference was designed based on multiple observation results corresponding to the classification results on the different brain regions using their corresponding SVM classifiers. The result of this inference was to classify the response pattern of the brain into either class 1 (relapser class) or class 2 (abstainer class).

For instance, we classify the response pattern of the brain for a particular unlabeled block i ($\dot{B}_i$) of an unseen test subject after the response patterns of K different individual ROIs ($\dot{p}_{ki}; k = 1 \ldots K$) extracted from the feature vector of the brain for the block i have been classified as $f_k(\dot{p}_{ki}) = y_{ki};\ k = 1 \ldots K;\ y_{ki} \in \{-1, +1\}$ using their corresponding SVM classifiers (Fig. 3.1 & 3.3). The inference process is conducted as follows.

Initially, when no observation has been provided, an initial classification probability of the response pattern of the brain for block i ($\dot{B}_i$) is set to 0.5 i.e. $P(\dot{B}_i = +1) = P(\dot{B}_i = -1) = 0.5$ (no difference to classify the pattern into either class 1 or class 2).

After observation y_{1i} from the first ROI ($f_1 = y_{1i};\ k = 1$) has been provided, Bayes' rule was applied to calculate the posterior probability that the pattern of the brain for block i ($\dot{B}_i$) was classified into class 1 (relapser class):

$$P(\dot{B}_i = +1 | f_1 = y_{1i}) = \frac{P(f_1 = y_{1i}|B = +1) \cdot P(\dot{B}_i = +1)}{P(f_1 = y_{1i})}$$

$$= \frac{P(f_1 = y_{1i}|B = +1) \cdot P(\dot{B}_i = +1)}{P(f_1 = y_{1i}|B = +1) \cdot P(\dot{B}_i = +1) + P(f_1 = y_{1i}|B = -1) \cdot P(\dot{B}_i = -1)}$$

in which, $P(\dot{B}_i = +1) = P(\dot{B}_i = -1) = 0.5$; $P(f_k = y_{1i}|B = +1)$ and $P(f_k = y_{1i}|B = -1)$ are the priors learned on the training data.

Now, the classification probability of the response pattern of the brain for block i is updated and used for the next inference: $P(\dot{B}_i = +1) = P(\dot{B}_i = +1|y_{1i})$; $P(\dot{B}_i = -1) = 1 - P(\dot{B}_i = +1)$.

Then, when the observation y_{2i} from the second ROI ($f_2 = y_{2i};\ k = 2$) has been provided as evidence, the posterior probability of the response pattern of the brain for block i ($\dot{B}_i$) can be calculated as

$$P(\dot{B}_i = +1 | f_2 = y_{2i}) = \frac{P(f_2 = y_{2i}|B = +1) \cdot P(\dot{B}_i = +1)}{P(f_2 = y_{2i}|B = +1) \cdot P(\dot{B}_i = +1) + P(f_2 = y_{2i}|B = -1) \cdot P(\dot{B}_i = -1)}$$

As before, the classification probability of the response pattern of the brain for block i ($\dot{B}_i$) is now updated: $P(\dot{B}_i = +1) = P(\dot{B}_i = +1|y_{2i})$; $P(\dot{B}_i = -1) = 1 - P(\dot{B}_i = +1)$.

The inference process continues until all relevant ROIs ($1 \dots K$) have provided their observation results (Fig. 3.3); and finally the response pattern of the brain for block i ($\dot{B}_i$) is classified into either class 1 (relapse class; $\hat{y}_i = +1$) or class 2 (abstainer class; $\hat{y}_i = -1$; $\hat{y}_i$: classification result of the response pattern of the brain for block i) according to the following decision rule:

$$\hat{y}_i = \begin{cases} +1 & if\ P\left(\dot{B}_i = +1 \middle| \{y_{ki}\}_{k=1..K}\right) \geq 0.5 \\ -1 & if\ P\left(\dot{B}_i = +1 \middle| \{y_{ki}\}_{k=1..K}\right) < 0.5 \end{cases}$$

(K: number of observed brain regions (ROI)

$i = 1, \dots, I_{testing}$; $I_{testing}$: number of the unlabeled patterns collected for the testing)

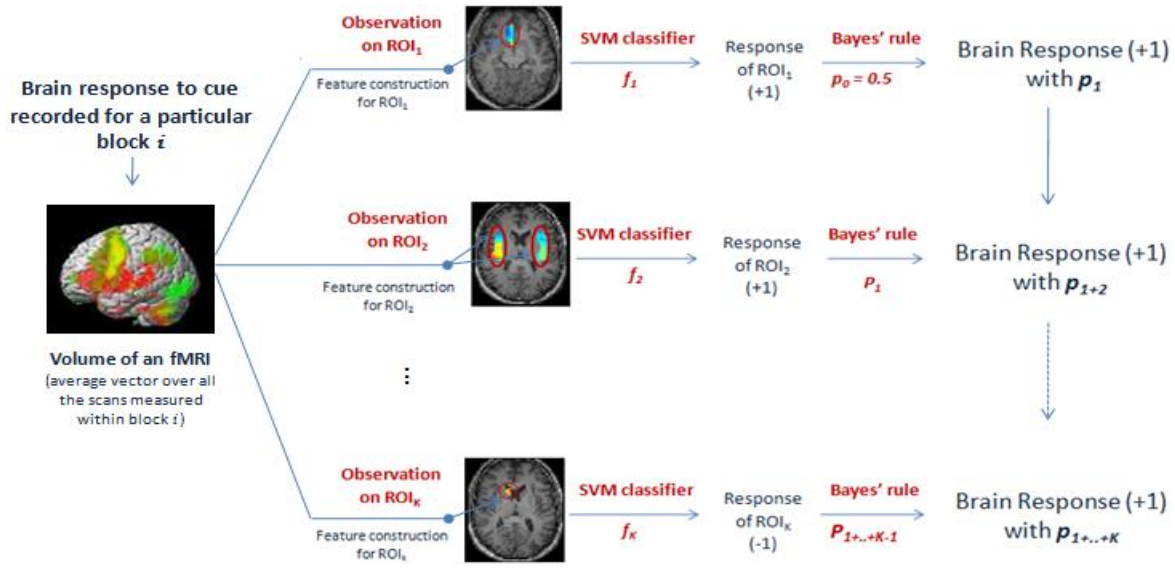


Figure 3.3. **Illustration for the inference to classify a response pattern of the brain recorded for a particular block i ($\boldsymbol{B_i}$) combining multiple observation results on multiple ROIs** (P: the probability to classify the pattern into class 1; and it is updated after each observation result on the ROI k has been provided $\left(P_{1+\dots+K} = P\left(\dot{B}_i = +1 \middle| \{y_{ki}\}_{k=1..K}\right)\right.$; y_{ki}: the observation result on the response pattern of ROI k recorded for block i).

This inference model was built on the basis of the model of independent relevant symptoms in which a disorder only involves some relevant symptoms, and the symptoms or the response features of ROIs are considered conditionally independent of each other. The priors used for the inference for each ROI k were estimated on the training set using an inner cross-validation loop (9-fold cross-validation procedure repeated five times) from the test results of the classification of response patterns of the ROI k with its corresponding SVM classifier f_k in step *A 1*. Specifically, the priors $P(f_k = +1|B = +1) = 1 - P(f_k = -1|B = +1) = p_{k+}$ and

$P(f_k = -1|B = -1) = 1 - P(f_k = +1|B = -1) = p_{k-}$ were estimated for each ROI k. The parameter p_{k+} corresponds to the sensitivity of the classifier f_k for the ROI k, while p_{k-} corresponds to the specificity of this classifier. Consequently, the priors can be considered as previously learned experience of the role of the ROI k in fMRI classification in the investigated disorder, and the inference based on such experiences and Bayes' rule can be considered as a way to emulate the way of decision-making in clinical practice (Kasper et. al., 2008).

Predictive inference of the response pattern of the brain

A predictive inference for a response pattern of the brain recorded for a block i can rely on either a single observation result (single evidence; K = 1) or a synthesis from multiple observation results on multiple ROIs (multiple evidences; K > 1). In order to evaluate the validity of the inference based on multiple lines of evidence in fMRI classification, the classification result of the response pattern of the brain based on a single observation on a single ROI were evaluated and compared with those based on multiple ROIs. For this, we applied the methodology just described. However, in this case, Bayesian inference for a response pattern of the brain for a particular unlabeled block i ($\dot{B}_i$) was only based on a single observation corresponding to the classification result of an SVM classifier (f_k) for an observed particular ROI k. As a result, the posteriors $P(\dot{B}_i = +1)$ and $P(\dot{B}_i = -1)$ used to infer block i ($\dot{B}_i$) after the observation result on the ROI k has been provided were calculated as follows:

$$P(\dot{B}_i = +1) = P(\dot{B}_i = +1|y_{ki}) = \frac{P(f_k = y_{ki}|B = +1)}{P(f_k = y_{ki}|B = +1) + P(f_k = y_{ki}|B = -1)}$$

$$P(\dot{B}_i = -1) = 1 - P(\dot{B}_i = +1)$$

for $k = 1, \dots, K$; K: number of observed brain regions

B. Classification of subject

Although the response patterns of the brain or brain regions is the main object that we would like to analyze in order to apply diagnostic imaging in practice, the final target is to predict disorder or condition of the disorder which subjects suffer from. In our context, this target is to classify subjects into given classes. For this, the classification of an unlabeled subject ($\dot{S}_a$) can rely on a synthesis of multiple observations on the various response patterns of the brain measured in

different blocks for the subject ($\dot{S}_a$). In the study, for each subject ($\dot{S}_a$), this synthesis was represented by the ratio of the number of response patterns of the brain that were classified into class 1 (relapser class) to the total number of the patterns measured for the subject ($\dot{S}_a$) ($R_{class\ 1} = \frac{\sum_{i=1}^{O}(\hat{y}_i=+1)}{O\ observations}$; $\sum_{i=1}^{O}(\hat{y}_i = +1)$: the number of observations where the response pattern of the brain for block B_i was classified into class 1; O: the total number of observed patterns of the brain (or number of observed blocks) for the subject $\dot{S}_a$). If this ratio was equal or greater than 0.5 ($R_{class\ 1} \geq 0.5$), the subject ($\dot{S}_a$) was classified as a relapser, and by contrast, if the ratio was smaller than 0.5 ($R_{class\ 1} < 0.5$), the subject ($\dot{S}_a$) was classified as an abstainer.

Evaluation

Creating examples for evaluation and learning

The creation of examples for evaluating and learning the hyper-parameters for SVM classifiers together with the priors for Bayesian inference in this study was done in a manner similar to that applied in the first one. As a result, a stratified 10-fold cross-validation procedure repeated 10 times (10-repeated 10-fold cross-validation procedure) was applied to create 10 rounds for the evaluation. In each round, the sample including 40 alcoholic patients in the study was partitioned randomly into 10 equal folds, each of which includes 4 alcoholic patients (2 relapsers and 2 abstainers, corresponding to 48 response patterns of the brain collected from them). Based on the 10 folds, each round created 10 loops of training and testing (L^t; t=1→10) in such a way that within each loop (L^t) a different fold of the data containing 4 patients (A^t) was held-out for testing while the remaining 9 folds, including the 36 other patients (B^t), were used for training. This procedure was repeated 10 times on random cross-validation splits of the 40-subject sample (L^t; t = 1 → 10 x 10). Moreover, within the training dataset (B^t) of each loop (L^t), a 9-fold cross-validation procedure repeated five times was applied to create 45 nested loops of training and testing ($nested\_L_j^t$; j = 5 x 9) for adjusting the hyper-parameters of the SVM classifiers used in step *A 1.2* and for learning the priors for Bayesian inference used in step *A 2 (*Fig. 3.1).

Evaluation

The evaluation and comparison of classification performance for each individual classifier for each ROI and between different classifiers were applied similarly to that used in the first study.

Results

A. Classification performance of the pattern of the brain

The results were summarized from testing 100 sample sets that were created from 10 repeated 10-fold cross-validations. Each set included 4 subjects for testing. Each subject comprised 12 response patterns of the brain or ROIs collected from 12 cue blocks. As a result, the total of the classified patterns for each observation on each ROI was 4800 (N = 4 x 12 x 100; Tables 3.3, 3.4 and 3.5). The indices of performance were the average values over all the sets reported (in %). Similarly, for evaluation of the significance of classifiers in the observations, their accuracies were said to be significant with 1% level under the Bonferroni correction (p_{corr}<0.01) if their uncorrected p-values were smaller than 0.425 x 10^{-7} (0.01/(Nx*m*); N: number of trials (classified patterns); N=4800; *m*: number of observed classifiers in the same context; *m* = 49; Tables 3.3, 3.4 and 3.5). According to the cumulative probability function on a binomial distribution, this also means that the accuracy of a classifier is significant with p_{corr}<0.01 if it is greater than 53.9% (uncorrected p-value = 0.306 x 10^{-7}; N = 4800; p = 0.5; p: probability of a success).

A.1 Classification performance based on a single evidence from an individual ROI

*1.1.1 Bilateral ROIs**

The results obtained from the classification of the response patterns recorded on individual bilateral ROIs are shown in Table 3.3 for the two different methods of normalizing features of the patterns (voxel attribute) before the classification (scaling normalization vs. the combination of z-score and scaling normalization). The classifiers on the VS and insula yielded the best performance with accuracies of 63.7% and 71.5% respectively (p_{corr}<0.0001). For several ROIs, the performance combining the z-score and scaling normalization was poorer than using the scaling normalization alone e.g. the VS, 52.7% vs. 63.7% accuracy respectively, χ_M^2 = 146.78, *p*<0.0001. In contrast, on the other ROIs, the combination of z-score and scaling normalization outperformed the scaling normalization alone such as the insula, 71.5% vs. 61% accuracy, χ_M^2 = 260.47, *p*<0.0001; the thalamus, 59.7% vs. 54.8% accuracy respectively, χ_M^2 = 61.23, *p*<0.0001 (Table 3.3).

Table 3.3. **The performance of pattern classification (%) for bilateral ROIs**

ROI N = 4800	Scaling normalization				Z-score and Scaling normalization			
	Code	*Accuracy*	*Sensitivity*	*Specificity*	*Code*	*Accuracy*	*Sensitivity*	*Specificity*
VTA	1	55	52.1	57.9	12	33.3	46.6	20
VS	2	**63.7**	68.6	58.8	13	**52.7**	52.1	53.3
DS	3	38.8	44.4	33.2	14	40.7	28.2	53.2
Pallidum	4	37.5	33.5	41.5	15	39.8	36.6	43
Thalamus	5	54.8	55.3	54.3	16	59.7	62.9	56.4
ACC	6	49.9	46	53.8	17	54.8	55.8	53.8
mPFC	7	41.7	44.9	38.4	18	43.9	18.9	69
OFC	8	49.6	47.9	51.3	19	47.3	39.3	55.3
Amygdala	9	55.8	57.7	53.8	20	60.8	65.7	55.8
Insula	10	**61**	60.4	61.6	21	**71.5**	74.4	68.6
Hippocampus	11	47.1	47.5	46.7	22	49.4	50.8	48

**Note: bilateral ROI was abbreviated to name of the ROI e.g. bilateral insula was abbreviated to insula.*

1.1.2 Separate ROIs for left and right hemisphere

In order to check for possible lateralization effects, we repeated the analysis conducted on the bilateral ROIs in section 1.1.1 on separate ROIs of the left and right hemispheres. For simplification, we applied the combination of z-score and scaling normalization in the feature construction of all the ROIs, except the VTA and VS. The results are reported in Table 3.4. This analysis showed that there existed an asymmetry in the classification performance on several ROIs. Especially, the performance on the right structure of VS and ACC was significantly higher than the performance on the left of VS and ACC: VS (75.9% vs. 53.1% accuracy; χ^2_M = 603.09) and ACC (68.2% vs. 58.9%; χ^2_M = 145.94); p<0.0001; Table 3.4).

Table 3.4. **The performance of pattern classification (%) for the left and right ROIs**

ROI N = 4800	Left Hemisphere				Right Hemisphere			
	Code	*Accuracy*	*Sensitivity*	*Specificity*	*Code*	*Accuracy*	*Sensitivity*	*Specificity*
VTA	23	54.4	47	61.7	34	51.1	47.2	55
VS	24	**53.1**	48.4	57.8	35	**75.9**	73	78.9
DS	25	51.8	56	47.5	36	59.2	53.4	65
Pallidum	26	44.6	65.2	24	37	38.7	20	57.8

Thalamus	27	52.7	56.3	49.2	38	54.5	59.3	49.7
ACC	28	**58.9**	59.6	58.2	39	**68.2**	67.9	68.5
mPFC	29	52.2	51.8	52.6	40	42	49.9	34.1
OFC	30	45.5	34.4	56.5	41	57.5	58.8	56.1
Amygdala	31	52.5	48.5	56.5	42	53.3	34	72.5
Insula	32	50.8	48.8	52.8	43	44.1	39.7	48.5
Hippocampus	33	56.6	48.2	65	44	34.5	46.5	22.6

A.2 Classification performance combining multiple observation results on multiple ROIs

We used Bayesian inference to combine the predictions from the classifiers on individual ROIs. We constructed a predictor that included the three best performing individual ROIs (right VS, right ACC, and insula). Also, we investigated the inclusion of other well-performing regions (right DS, right OFC, amygdala, thalamus). The results are shown in Table 3.5. The combination of the right VS, right ACC and bilateral insula yielded a significantly higher accuracy than the individual ROIs (76.9% vs. the VS, 75.9% accuracy (χ^2_M = 3.94; p=0.0472); the ACC, 68.2% (χ^2_M=141.91; p<0.0001) and the insula, 71.5% (χ^2_M = 69.91; p<0.0001) respectively). However, the additional inclusion of another ROI, e.g. either the right DS or right OFC or amygdale, did not further improve the performance (76.9% vs. 73.8%, 74.4%, 72.2% accuracy; χ^2_M = 96.57, 110.25, 211.46 respectively; p<0.0001).

Table 3.5. **The performance of pattern classification by combining predictions on multiple ROIs**

ROI **N = 4800**	**ROI code**	**Classification performance (%)**		
		Accuracy	*Sensitivity*	*Specificity*
right VS - right ACC - insula	35 – 39 -21	**76.9**	76.3	77.5
right VS - right ACC - insula - right DS	35 – 39 -21 - 36	73.8	74.3	73.3
right VS - right ACC – insula – right OFC	35 – 39 -21 – 41	74.4	74	74.8
right VS - right ACC - insula - amygdala	35 – 39 -21 - 20	72.2	72.5	71.8
right VS - right ACC - insula - thalamus	35 – 39 -21 - 17	73.9	73.6	74.2

B. Classification of subject

As mentioned in the section on methodology, the classification of a subject as either a relapser or an abstainer was designed to rely on the synthesis of observation results on the response patterns

of the brain measured in various blocks for that subject, and the classification results can be obtained from the synthesis of various observations either on a single ROI or on multiple ROIs. The results of the two methods are shown in Table 3.6. The accuracies of subject classification that were said to be significant with 1% level under the Bonferroni correction were the ones whose p-values were smaller than 0.208×10^{-5} (0.01/(N×*m*); N: number of trials (classified subjects); N = 400; *m*: number of observed classifiers in the same context; *m* = 12; Table 3.6). This also means that the accuracy of a classifier is significant with p_{corr}<0.01 if it is greater than 61.4% (uncorrected p-value = 0.15×10^{-5}; N = 400; p = 0.5; p: probability of success in one trial).

The results showed that there was no significant difference between the classification of the response pattern of the brain ('block classification') and the classification of subject (p>0.05). The classification performance of a combination of the right VS, right ACC and insula was better than that of the respective individual ROIs (77% vs. the right VS, 76.5% accuracy (χ^2_M = 0.02; p=0.8852); the right ACC, 68% (χ^2_M = 12.01; p=0.00053); the insula, 71.3% (χ^2_M = 5.69; p=0.017). In the other combinations of the right VS, right ACC, insula and either the right DS or right OFC or amygdala, the overall performance was not better than the performance of a combination of the right VS, right ACC and insula (73.8%, 74.5%, 72% vs. 77% accuracy; χ^2_M = 7.58, 8.1, 18.05; p=0.0059, 0.0044, 0.00002 respectively; Table 3.6).

Table 3.6. **The performance of subject classification**

ROI N = 400	ROI code	Classification performance (%)		
		Accuracy	*Sensitivity*	*Specificity*
right VS	35	76.5	73	80
right DS	36	57.5	54	61
right ACC	39	68	68	68
right OFC	41	57.8	59.5	56
Insula	21	71.3	73.5	69
Thalamus	17	60.5	64.5	56.5
Amygdala	20	61	68	54
right VS - right ACC - insula	35 – 39 -21	**77**	76	78
right VS - right ACC - insula - right DS	35 – 39 – 21 - 36	73.8	74.5	73
right VS - right ACC - insula - right OFC	35 – 39 – 21 - 41	74.5	73.5	75.5
right VS - right ACC - insula - amygdala	35 – 39 – 21 - 20	72	72.5	71.5
right VS - right ACC - insula - thalamus	35 – 39 – 21 - 17	74	73.5	74.5

C. Classification of the response pattern of the brain in the cases where the patterns of combined ROIs have been classified into the same class (either class 1 or class 2)

As mentioned above, the classification of a response pattern of the brain was designed to be based on the combination of observation results on the response patterns of several relevant ROIs extracted from its feature vector (Fig. 3.1). As a result, there may be a number of samples (response patterns of the brain) (column N_1 of Table 3.7) in which the classification results of the response patterns of combined ROIs were in the same class e.g. $f_{1i}(x_{1i}) = f_{ki}(x_{ki}) = f_{Ki}(x_{Ki}) = +1$ or $f_{1i}(x_{1i}) = f_{ki}(x_{ki}) = f_{Ki}(x_{Ki}) = -1; 1 \dots K$: combined ROIs; i: the block from which the feature vectors of ROIs $1 \dots K$ were extracted. By contrast, there may be also species of samples (response patterns of the brain) (column N_2 of Table 3.7) in which the classification results of the response patterns of combined ROIs disagreed e.g.$(f_{1i}(x_{1i}) = +1) \neq (f_{ki}(x_{ki}) = -1)$ (or right VS (f = +1) and right ACC (f = -1)).

The investigation on all the samples where the classification results of response patterns of combined ROIs extracted from the same feature vector of the brain were in the same class showed that if the response patterns of 2 ROIs (e.g. right VS-right ACC or right VS-insula or right ACC-insula) were classified into the same class, the classification accuracy recorded on the focused sample set was up to 86% (Table 3.7); and if the response patterns of these 3 ROIs were classified into the same class, the classification accuracy was over 96% (Table 3.7). Especially, once the investigation was focused on the samples where the response patterns of 4 ROIs of the right VS, right ACC, insula and right DS were classified into the same class, the classification accuracy achieved 98.9%.

Table 3.7. **Classification performance of the response patterns of the brain in the cases where the response patterns of combined ROIs have been classified into the same class** (N: total number of samples or response patterns of the brain created from 10 repeated 10-fold cross-validation procedure on 40 subjects (N = 4800); N_1: number of samples (or the response patterns of the brain) which have the agreement classification results of response patterns of combined ROIs extracted from the same feature vector of the whole brain; N_2: number of samples (or the response patterns of the brain) which have the disagreement classification results of response patterns of combined ROIs extracted from the same feature vector of the whole brain).

ROI	N = 4800		Classification performance on N_1 (%)		
	Percentage of excluded samples N_2/N	Percentage of analyzed samples N_1/N	*Accuracy*	*Sensitivity*	*Specificity*
right VS	0	100	75.9	73	78.9

right ACC	0	100	68.2	67.9	68.5
Insula	0	100	71.5	74.4	68.6
right VS - right ACC	36	64	86.7	84.4	88.8
right VS - insula	37.75	62.25	86	84.4	87.6
Insula - right ACC	43.25	56.75	86.8	89.1	84.6
right VS - right ACC - insula	58.65	41.35	96.2	97.9	94.6
right VS - right ACC - insula – right DS	76.56	23.44	98.9	97.6	99.8
right VS - right ACC – insula – right OFC	74.92	25.08	95.9	99.6	92.7
right VS - right ACC - insula – amygdala	78.56	21.44	98.7	98.8	98.6
right VS - right ACC - insula – thalamus	71.1	28.9	95.8	97.9	94.2

Discussion

The results in this study provide further evidence that fMRI can identify biomarkers to predict relapse after detoxification in alcohol dependence (Heinz et al., 2009; Beck et al., 2012). Besides, our approach where the activation of relevant brain regions is observed separately and then the observation results are combined appropriately can have benefits in the investigation of the role of the brain regions of interest as well as the method of predictive inference based on multiple lines of evidence in the analyzed disorder. On the other hand, the results showed no significant difference between the classification results of the response pattern of the brain and the classification results of subject (p>0.05; Tables 3.3 & 3.6). The data can be taken as evidence for the validity of predictive inference of a neuropsychiatric disorder or its condition using a synthesis of multiple observations of activation feature of the brain measured by fMRI whenever the brain is stimulated. Such an approach is compatible with diagnostic imaging and leaves open the opportunity to analyze temporal characteristics of activation data in further studies.

Investigation on individual ROIs

Insula in relapse prediction

With a high accuracy of 71.5% (p_{corr}<0.0001; Table 3.3), the insula has demonstrated its important role in the underlying mechanism mediating relapse. The result is consistent with recent evidence indicating that the insula plays a crucial part in conscious urges to take drugs, which can precipitate relapse (Craig, 2009; Naqvi & Bechara, 2009; Vincent et al., 2012). The insula has been known as a region that integrates interoceptive states into conscious feelings and

into decision-making processes involving uncertain risk and reward (Naqvi & Bechara, 2009). Research results on smokers who sustained damage in the insula showed that these patients quit smoking and remained abstinent much more easily and immediately than smokers who sustained damage in other brain areas (Naqvi et al., 2007). In addition, a number of functional imaging studies have shown that the insular cortex is activated when drug abusers are exposed to drug cues that trigger craving (Tapert et al., 2004; Filbey et al., 2009; Vincent et al., 2012). A study of response to conditioned alcohol taste cues following experience with the drug in rats indicated that ethanol intake and preference were formed during the chronic exposure phase, and inactivation of the insula eliminated this preference (Castro, 2012). This evidence suggests that the insula may play an important role in the mechanism leading to relapse and may be one of the major targets for the treatment of drug addiction.

Lateralization

The investigation results of lateralization indicate that the structures of VS and ACC in the right hemisphere contain much more relevant information distinguishing relapsers from abstainers than those in the left one. In other words, a functional asymmetry appears to exist in response to alcohol-associated cues between the two hemispheres in the striatum and ACC. This result is in line with results of previous studies indicating that the right hemisphere is more vulnerable to the effects of alcohol than the left (Oscar-Berman, 2003). The effects can be characterized by a decreased volume of structures involving reward system, which occur in the right hemisphere more pronouncedly than those in the left (Makris et al., 2008), and by the response feature to alcohol-associated cues showing that activation is more stable in the right VS than the left (Schacht et al., 2011). On the other hand, the observations of Makris et al. (2008) showed that the volume of VS increased with length of abstinence in alcohol-dependent patients, confirming that brain atrophy can be partly reversible, and the recovery of such damage may be a predictor for abstinence (Bühler & Mann, 2011). Taken together, the results of the study provide evidence for the existence of lateralization for relapse prediction in the brain regions of VS and ACC in response to alcohol cues. However, the source for the asymmetry still needs further study because there is a paucity of data on functional lateralization in relapse prediction. Recently, a study of Beck et al. (2012) showed an increased brain response of abstainers to cues in both of the left and right VS compared with the response of relapsers. The result supports for our result

indicating that both of the left and right VS contain valuable information for the classification. However, we here found that there was a lateralization for relapse prediction in this structure where the response patterns of VS recorded in the right hemisphere were more predictive than those in the left hemisphere (the left VS, 53.1% vs. the right VS, 75.9% accuracy; χ^2_M = 603.09; p<00001). The method of analysis in the study of Beck et al. (2012) was to identify significant activation voxels in the VS distinguishing relapsers from abstainers, obtained from a statistical analysis using a univariate approach. This approach is different from our approach which is to classify the individual activation patterns into the two groups using a multivariate method based on all the voxels of the investigated structure. Moreover, previous studies of fMRI classification show that there exists a difference in feature selection between the univariate and multivariate approaches (Mourão-Miranda et al., 2005), and the multivariate approach fared better than the univariate approach in identifying category information in investigated structures (Kriegeskorte et al., 2006; Normand et al., 2006; Pereira et al., 2009). In any case, the evidence indicates that the response feature of the VS to alcohol-associated cues is an important indicator for prognosticating relapse, which was confirmed in our study.

Validity of deeper focus on structural ROIs

The classification performance of bilateral insula with a size of 1215 voxels was significantly higher than the performance of both the left insula with a size of 612 voxels and the right insula with a size of 603 voxels (71.5% vs. 50.8% and 44.1% accuracy; χ^2_M = 537.83 and 790.81 respectively; p<0.0001; Tables 3.3 & 3.4). In contrast, the performance of the right VS with a size of 44 voxels and the right ACC with a size of 489 voxels were significantly higher than the performance of the bilateral VS with a size of 88 voxels and the bilateral ACC with a size of 987 voxels (75.9% vs. 63.7% accuracy (χ^2_M = 383.71) and 68.2% vs. 54.8% (χ^2_M = 337.34) respectively; p<0.0001; Tables 3.3 & 3.4). The results reinforce the assumption that the performance of a ROI in relapse prediction may not depend on its size but on its specific characteristics in response to given stimulation. On the other hand, as discussed in the first study, since the nature of multivariate methods such as SVM is to exploit attributes of all input variables (voxel attribute) to categorize (Wang, 2005; Vapnik, 2000), a classification decision for a pattern of a ROI using this method relies not only on individual voxels but also on correlations between the voxels within that ROI. The presence of too many voxels which are insignificant for

classification or do not correlate closely with the other significant voxels within the investigated ROI can deflect the dominant classification direction of the significant voxels and cause misclassification (Pereira et al., 2009). In such a situation, homogeneous attributes of anatomical structure and function can bring advantages for multivariate analysis, since local connections between neurons of similar specificity forming neuronal groups are stronger and more complex than those between neurons of the various functional groups (Tononi, 1994) or in other words, the features of individual voxels or input variables within the investigated structure are then placed in close correlation. As a result, the low performance of the other structures e.g. the bilateral OFC (49.6% accuracy) achieved in this study may originate from our method of feature selection, which can be improved if their investigated region is localized more homogeneously and exactly in terms of the structure and function (e.g. the bilateral OFC 49.6% vs. the right OFC 57.5% accuracy; χ_M^2 = 106.06; p<0.0001; Tables 3.3 & 3.4). On the other hand, due to difference of location, structure and function that can impact on fMRI measurements (Mathews et al., 2001), each ROI can require compatible methods of analysis such that it can expose its distinguishable response features in the best possible way. This is evidenced by the significant differences of classification performance on several ROIs between using the combination method of z-score and scaling normalization and the single method of scaling normalization such as the VS, 52.7% vs. 63.7% accuracy respectively, χ_M^2 = 146.78, p<0.0001; the insula, 71.5% vs. 61% accuracy respectively, χ_M^2 = 260.47; p<0.0001 (Table 3.3).

Validity of combining multiple observation results on multiple ROIs

Given the situation that there is no ROI playing a decisive role in prediction as a "gold standard", the prediction can be based on an extensive observation on multiple relevant ROIs to obtain higher accuracy compared with predictions only based on a single observation on a single ROI. The assumption is supported by the better performance combining multiple observation results on the VS, ACC and insula than those using a single observation result on an individual ROI (76.9% vs. the right VS, 75.9%, the ACC, 68.2% and the insula, 71.5% accuracy respectively; p<0.05; Tables 3.3 & 3.5). The result provides evidence that specific response features of the brain to specific stimuli can spread across several relevant brain regions (Barry et al., 1999), and that they can be identified and integrated into an overall picture used for predictive inference of brain disorder. This also suggests that the response patterns of different brain regions with

different structures and functions can be observed independently, and the more the brain regions manifest the response features that are classified into the same class, the higher the classification probability into that class (Table 3.7). Such an approach can enable us to prognosticate relapse risk via functional imaging with an estimable confidence. Additionally, our study was conducted on the sample with no significant difference in clinical data between the two groups of subsequent relapsers and abstainers after detoxification (Table 3.1), and fMRI measurements were carried out about 2 weeks after detoxification, before participants were identified as relapsers or abstainers 6 months later (Table 3.1). These data provide evidence demonstrating that fMRI can identify biomarkers for relapse prediction in alcohol dependence.

An addition of ROIs such as either the right DS with 59.2% accuracy or the right OFC with 57.5% accuracy (or the amygdala with 60.8% accuracy) to the combination of the right VS, right ACC and insula did not yield better performance (73.8%, 74.4%, 72.2% vs. 76.9% accuracy respectively; $p<0.0001$; Tables 3.3-3.5). However, if the observation was only focused on the samples where the response patterns of all the combined ROIs extracted from the same feature vector of the brain were classified into the same class, the classification accuracy on this sample set increased linearly with the number of combined ROIs from appropriately 86% accuracy for the combination of two of the three ROIs: right VS, right ACC and insula to 98.9% accuracy for the combination of four ROIs: right VS-right ACC-insula-right DS (Table 3.7). Unfortunately, unlike the accuracy increase, the number of agreement samples diminished correspondingly from appropriately 60% to 23.44% (Table 3.7). This shows that there existed greater differences of classification results of the response patterns of individual ROIs when the number of the combined ROIs was increased, and misclassification of the added ROIs (e.g. right DS) can deflect the correct inference direction of the other ROIs in some analyzed samples via Bayesian inference (Stefan & Lionel, 2011) and lower the overall performance. Despite this, in a positive aspect, the results reinforce the potential for increasing classification accuracy by methods of controlling the combination of the evidence collected from the ROIs in the inference process. One of the commonly used methods in clinical practice is to eliminate from the inference process ambiguous signs or symptoms that in this context are the indistinguishable response patterns of ROIs which can cause misclassification. How to identify the indistinguishable response patterns reliably and whether it may be a feasible approach to diagnostic functional imaging in practice are the issues which will be addressed in the next chapter.

CHAPTER IV

IMAGING APPROACH IN fMRI CLASSIFICATION

Introduction

In the first and second studies, the classification of response patterns of the investigated ROIs was implemented automatically by machine. In order to realize diagnostic functional imaging in clinical practice, our third study was designed for the initial step of imaging approach based on the indication of machine-based image recognition. Specifically, it offers a way to allow us to have insights into individual response patterns accompanied with their corresponding classification decisions obtained from machine, by which means we can gradually learn and find out rules of the classification for the response patterns of the investigated brain regions. This would open an important window into differences of individual functional images identified by machine inference between the two investigated classes, which would allow moving beyond group-based analysis and to the important clinical realm of diagnostic imaging.

Materials and Methods

Materials

This study was a continuation of the second study, which delves deeper into the investigation of machine-based classification decisions and discernibility level of classified patterns between the two classes, i.e. relapsers and abstainers. Thus, the materials and data pre-processing in this study were those used in the second study.

Methods

Outline of the method

The response patterns of each brain region collected from relapsers and abstainers were considered as experience of response imaging of that region to given stimulation for the two classes. To learn the experience better, these patterns should be ranked according to different level of the response feature between the two classes. The ranking would form a data bank of the ranked response patterns for each brain region for relapse risk which can help to obtain an insight

into the response images of this brain region more easily and to facilitate classification. The method was designed in three steps: *Step 1:* constructing and collecting the response patterns of ROIs; *Step 2:* ranking the collected patterns; *Step 3:* validating the ranking.

Step 1: Constructing and collecting the response patterns

The feature selection for individual ROIs was applied in a manner similar to that used in the second study. Consequently, for a particular ROI k, each subject provided 12 response patterns ($p_{ki};\ i = 1 \rightarrow 12$); and for 20 relapsers and 20 abstainers, 480 response patterns were collected and interpreted as 480 independent observations of response images of the ROI k to alcohol cues. They were used as input data of the corresponding classifier f_k.

Based on the findings in the previous studies of relapse in alcohol dependence as well as the results in the second study (Makris et al., 2008; Schacht et al., 2011; Beck et al., 2012), the brain regions, which play an important role in relapse prediction, were selected for the investigation in this study including the right VS, right ACC and insula. As those in the second study, the functional ROIs for these brain regions were also just their corresponding structural ROIs.

Step 2: Ranking the response patterns of individual ROIs

Defining index of the ranking & algorithm of the ranking

As mentioned above, the set of the 480 response patterns for each ROI k may be viewed as experience or data bank of response imaging of that ROI of alcoholic patients to the given stimulation accumulated from the two investigated groups comprising 20 relapsers (class 1) and 20 abstainers (class 2). The ranking was carried out based on the assumption that in the collected patterns there may be the patterns showing a prominent difference between the two classes, while the other patterns manifest the confused or indistinguishable difference and that the different level can be detected through classification. From this assumption, the classification result of a pattern (p_{ki}) can be used to measure the difference level of that pattern. This difference level was also defined as the discernibility level showing the degree of separation of the pattern between the two classes (Voulgaris & Mirkin, 2008). Moreover, from validity standpoint, it is logical that the measurement obtained from a single classification may be less reliable than those obtained from a synthesis on multiple classifications (Kukar & Kononenko,

2002; Voulgaris & Mirkin, 2008), and the degree of agreement of the classification decisions over the various classifications would give an indication of consistency of the discernibility level of that pattern (Douglas, 2007).

On this basis, for each ROI k, the ranking index of a response pattern (p_{ki}) for a class (c_a) was defined as the 'classified index' of the pattern (p_{ki}) into class (c_a) which was estimated by the ratio of the number of tests in which the pattern (p_{ki}) is classified into the class (c_a) to the total number of tests. In the study, a pattern (p_{ki}) was classified into one of the two classes, so the classification value of the pattern (p_{ki}) for class 1 is the complement of class 2. Likewise, the ranking index of the pattern (p_{ki}) for class 1 is the complement of class 2. For this reason, we only considered the ranking index (RI) of the pattern (p_{ki}) for class 1 (relapser class); it ranged from 1 to 0, with level '1' indicating the most discernible response pattern and level '0' indicating the lowest discernible response pattern for relapser class or in reverse order, indicating the most discernible pattern for abstainer class (class 2). Given this design, the ranking index (RI) of the pattern (p_{ki}) can be seen as the index of relapse risk for that pattern (p_{ki}). On this basis, we hope that for each ROI, a ranked scale of the response patterns with the ranking index of relapse risk would be formed which can facilitate imaging approach and which can be used in classification.

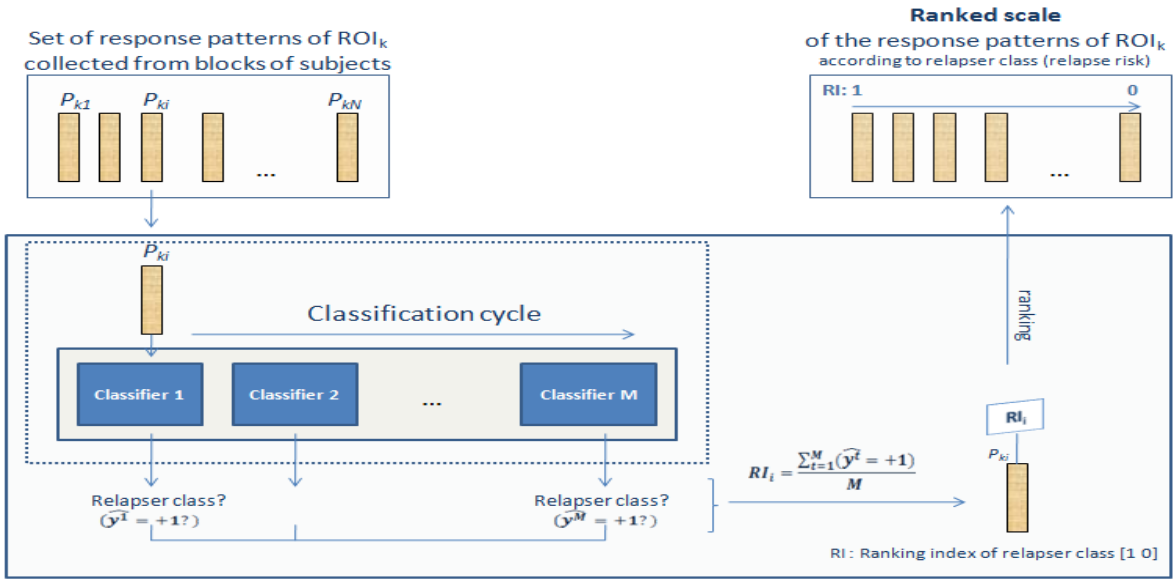


Figure 4.1. **Illustration for the ranking algorithm.** The ranking of a particular pattern (p_{ki}) is designed to translate the machine-based decisions of class classification for the pattern (p_{ki}) obtained from M classifications into the ranking index of relapse risk for that pattern.

Creating examples and calculating ranking index

The 10-repeated 10-fold cross-validation procedure was applied similarly to that used in the second study. However, in this study, for each loop of training and testing (L^t), instead of a single training set (B^t), each training segment ($nested\_B_j^t$) of each nested loop ($nested\_L_j^t; j =$ 5 x 9) within the training set B^t created from the stratified 9-fold cross-validation procedure repeated 5 times was used as a separate training dataset for the loop (L^t), and the corresponding testing set (A^t) was the dataset including the 4 patients used for the testing of the loop (L^t) (Fig. 4.2).

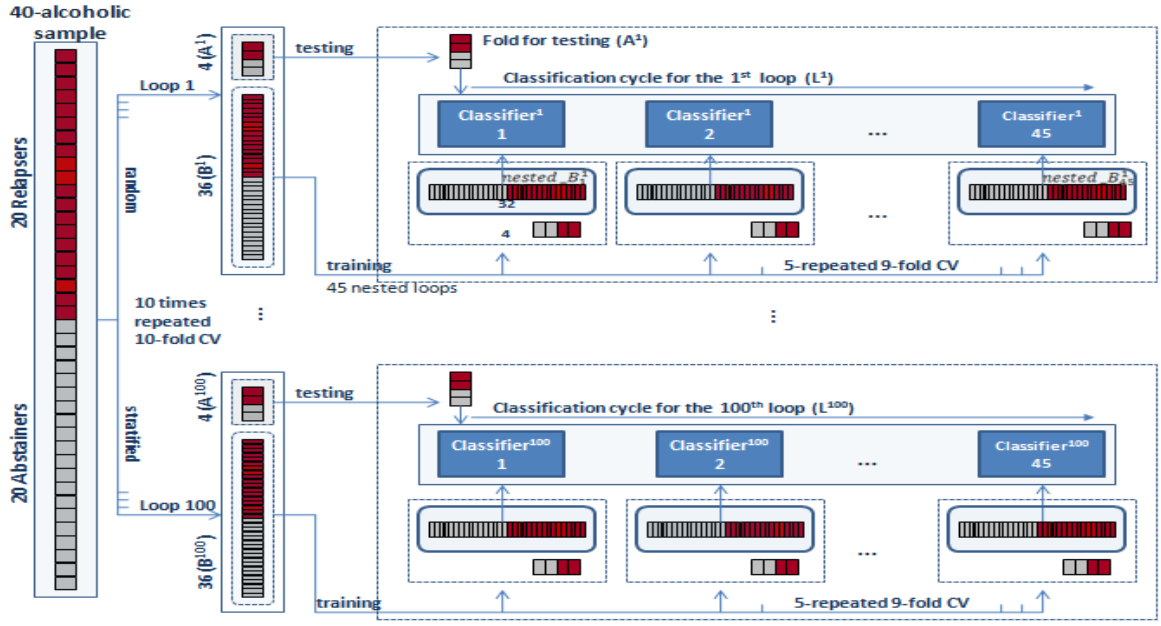


Figure 4.2. **Creating examples for calculating the ranking index of relapse risk** (CV: cross-validation).

With this design, each of the response patterns (p_{ki}) recorded for the 4 patients of the testing dataset (A^t) was classified 45 times for each loop (L^t). Based on these classification results, the ranking index for the pattern (p_{ki}) was estimated. To avoid an optimistic result (Pereira et al., 2008), the ranking index for each pattern (p_{ki}) was estimated only for the classifications where this pattern has been partitioned into the testing set (A^t). With the cross-validation procedure, for each round, a different fold of the data was used in turn for testing. This also means that for each round, a particular subject or the patterns (p_{ki}) recorded for that subject appeared only once in a testing segment. As a result, for 10 rounds of the 10-repeated 10-fold cross-validation procedure, each pattern (p_{ki}) was classified 450 times (M = 45 x 10).

Classifying the response patterns of ROI

a. Similarity and dissimilarity between the response patterns

With the method of feature construction described earlier, a response pattern (p_{ki}) of each ROI k was a feature vector holding the attributes of all the voxels of the ROI k recorded within a block (Fig. 4.3). Therefore, the similarity and dissimilarity between the two different patterns may be characterized by the distance between them with the rule that the larger the distance, the larger the dissimilarity between the two patterns and vice versa (Duda et al., 2001). There are several distance measures, one of which is estimated by the Euclidean distance (McCune & Grace, 2002). The formula of the Euclidean distance is applied as follows:

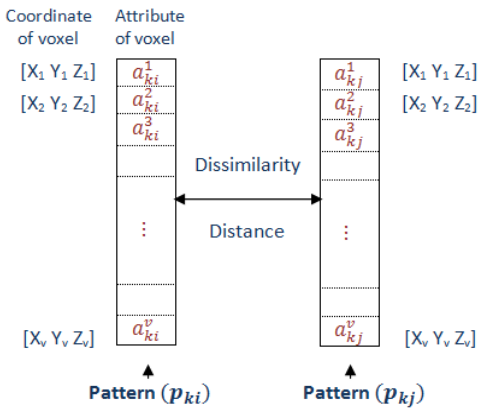


Figure 4.3. **Attribute of the response pattern**

$$d(p_{ki}, p_{kj}) = d(p_{kj}, p_{ki}) = \sqrt{(a_{kj}^1 - a_{ki}^1)^2 + (a_{kj}^2 - a_{ki}^2)^2 + \ \ldots + (a_{kj}^v - a_{ki}^v)^2} = \sqrt{\sum_{t=1}^{v} (a_{kj}^t - a_{ki}^t)^2}$$

Where p_{ki}, p_{kj} are two different response patterns; $d(p_{ki}, p_{kj})$ and $d(p_{kj}, p_{ki})$ are the Euclidean distance between the two patterns; (a_{kj}^1, a_{ki}^1), (a_{kj}^2, a_{ki}^2), … , (a_{kj}^v, a_{ki}^v) are the pairs of corresponding variables (or attributes of voxels) of the patterns p_{ki} and p_{kj}; [X_t Y_t Z_t] are the corresponding coordinates of voxels in the MNI standard space of investigated ROI (t=1→ v; v is the size of the ROI) (Fig. 4.3).

On this basis, a pattern can be simply classified into a class (c_a) if the dissimilarity of the pattern to that class (c_a) is the smallest, or in other words, the distance from the pattern to the representative pattern of the class (c_a) is the closest. The representative pattern of a class may be defined by the mean vector of that class (Schölkopf & Smola 2000; Balakrishnama & Ganapathiraju, 2013). This simple method of pattern classification may be appropriate if data for the classification are spread evenly in all directions, and the feature space is isotropic (Duda et al., 2001). However, this is the problem for original fMRI data because they are often very high dimensions in the feature space (Song et al., 2009). The obstacle has prompted the formation and development of algorithms of pattern classification for such data.

b. Classifiers

The algorithm was applied to pattern classification in this study in the same way as in the second study. Likewise, for each ROI k, a separate SVM classifier f_k was also used to classify its response pattern (p_{ki}) into either class 1 (relapse class) or class 2 (abstainer class). However, in this study, besides the class classification for the response pattern (p_{ki}) (into either class 1 or class 2), we considered the confidence level of the classification decision. As mentioned in the methodology in the first study, the class classification for a pattern (p_{ki}) relies on sign of the distance of the classified pattern (p_{ki}) from the optimal margin hyperplane of the classifier ($f_k(p_{ki}) = sign(w^T \emptyset(p_{ki}) + b)$; Fig. 2.3) while the confidence level of this decision depends on the absolute value of this distance ($\left|w^T \emptyset\left(p_{ki}\right) + b\right|$) since this value shows the degree of separation of the pattern (p_{ki}) between the two classes. Specifically, this implies that the larger the absolute decision value, the higher the confidence of the classification decision (Vlachos, 2004). Thus, this value can be used to measure the discernibility level of the response pattern. Nonetheless, it is an uncalibrated value mapped in R (Platt, 2000; Vapnik, 2000). Moreover, the calculation of this decision value depends strictly on the support vectors identified during the training. Accordingly, the decision values produced by SVM classifiers with different kernels or with different training datasets cannot be used to compare each other (Vapnik, 2000). In order to get the estimate more standardized, the decision values are mapped into probabilities. For a detailed description, reference is made to Platt, 2000; T.F. Wu et al., 2004. In the study, both of the two estimates were evaluated and compared with the ranking index. Moreover, to evaluate outperformance of classification with SVMs, the classification results of individual ROIs with this method were compared to those obtained with the method in which data were classified in their original input space with the simple decision rule based on the difference of the Euclidean distance from the mean vectors of the two classes.

The normalization of data as well as the selection of kernel and hyper-parameters for SVM classifiers was applied in a way similar to that used in the second study. Based on the classification performance of the ROIs in the second study both the scaling and z-score normalization were applied to the right ACC and insula while only the scaling normalization was applied to the right VS.

Step 3: Validating the ranking

Ranking index, decision value, and probability estimate

The ranking index of a pattern (p_{ki}) was designed as an indicator of the discernibility level of the pattern between the two investigated classes and was estimated by a synthesis of class decisions for the pattern from various classifications. Thus, the ranking index of a pattern was seen as the result of measurement for its nature. With such a design, the key to determining the validation of the ranking is validity and reliability of the measurement. Fortunately, for SVM classifiers, the discernibility level of response patterns between the two investigated classes can be characterized by their outputs such as the decision values or probability estimates (Platt, 2000; T.F. Wu et al., 2004; Vlachos, 2004; Voulgaris & Mirkin, 2008). From this basis, to investigate the validity of the ranking index corresponding to what it is supposed to measure (Kimberlin & Winterstein, 2008), the ranking index of the response pattern for each ROI k was compared with the corresponding decision value and probability estimate through correlation between them, and for evaluation of the reliability, we investigated stability and consistency through convergence of the index (Kimberlin & Winterstein, 2008).

For investigation of the correlation, the ranking index of each pattern obtained after M classifications was calculated according to the definition described earlier while the decision value and probability of that pattern was the average value of all the corresponding decision values and probabilities over the M classifications (M = 450).

For the convergence, we investigated variation of the ranking index of the response patterns during M classifications. For each ROI k there were the collected 480 patterns, each of which was classified 450 times. Consequently, a representative index for the 480-pattern set at the j^{th} classification was estimated by an average value (aRI^j) over the 480 ranking indices of the 480 patterns obtained after j classifications (j=1 → M; M = 450; Fig. 4.4). Together with the observation of the variation of this representative index during M classifications, its corresponding expectation value ($E(aRI^j) = \frac{\sum_{t=1\rightarrow j} aRI^t}{j}; j = 1 \rightarrow M; M = 450$) and error rate ($Er^j$) compared with the expected ranking index at the last classification ($Er^j = \frac{|aRI^j - E(aRI^M)|}{E(aRI^M)}; j = 1 \rightarrow M; M$ = 450) were taken into account as well.

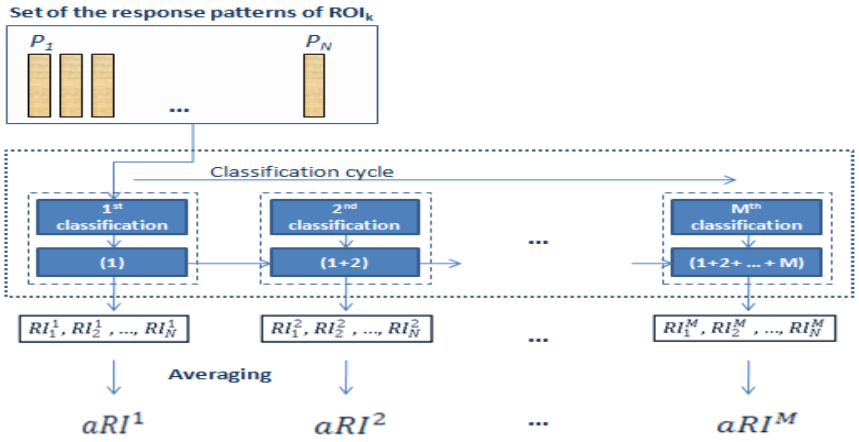


Figure 4.4. **Investigation of convergence of the ranking index** (N = 480 for each ROI; M = 450; RI_i^j: the ranking index of the pattern i at the j^{th} classification; aRI^j: the averaging ranking index over all the patterns (N = 480) at the j^{th} classification).

Ranking index and classification

a. Classification of the response patterns of individual ROIs

After ranking, the 480-pattern set for each ROI k can be viewed as the scale of the unlabeled response patterns of that ROI ($\dot{p}_{ki}$) for relapse risk accumulated from 20 relapsers (class 1) and 20 abstainers (class 2). The ranking index of a pattern ($\dot{p}_{ki}$) was considered as the index showing the position of that pattern on the scale. Hence, the ranking index is only an indicator of the discernibility level of the response pattern for relapse risk rather than a label of the pattern ($\dot{p}_{ki}$) in either class 1 or class 2. However, thanks to this indication, it can be used to classify the pattern ($\dot{p}_{ki}$) with the rule that if the ranking index of the pattern ($\dot{p}_{ki}$) is equal or greater than 0.5 (RI≥0.5), the pattern is classified into class 1, and by contrast, if the index is smaller than 0.5 (RI<0.5), the pattern is classified into class 2. In the study, the application of the ranking index in classification was carried out at the last classification (M = 450), and for each ROI, the performance was evaluated on the classification results of all the response patterns of that ROI (N = 480).

To compare the ranking based on the decision value and probability estimate, the classification results of the response patterns based on the average decision value ($aDec$) and the average probability estimate ($aProb$) summarized from the M classifications were considered as well (M = 450). The decision rule for these classifications was similar to those made by SVM classifiers. For the decision value, if the average decision value ($aDec$) of a pattern ($\dot{p}_{ki}$) is equal or greater than 0 ($aDec \geq 0$), the pattern is classified into class 1, and by contrast, if $aDec < 0$, the pattern is classified into class 2. For the probability estimate, if the average decision value ($aProb$) of a

pattern ($\dot{p}_{ki}$) is equal or greater than 0.5 ($aProb \geq 0.5$), the pattern is classified into class 1, and conversely, if $aProb < 0.5$, the pattern is classified into class 2.

b. Classification of subject

As in the second study, the classification of a unlabeled subject ($\dot{S}_a$) as either a relapser or an abstainer was also designed based on a synthesis of multiple observation results on the response patterns of the brain measured in various blocks for the subject ($\dot{S}_a$). However, instead of the synthesis based on the observation results of class on these patterns in the second study, in this study the synthesis was designed to specify the ranking index of relapse risk for the response feature of each brain region of interest for the subject ($\dot{S}_a$). For this, the ranking index for the response feature of the ROI k for each subject (sRI^k) was calculated by averaging all the ranking indices of the response patterns of the ROI k measured in the corresponding blocks for that subject ($sRI^k = \frac{\sum_{i=1}^{O} RI_i^k}{O\ observations}$; RI_i^k: ranking index for the patterns of ROI k recorded in block i; O: number of blocks for a subject; O = 12). Then, the rule for the classification of subject was applied in a manner similar to that used for the classification of pattern just described above. Specifically, in the case of the classification of a particular subject ($\dot{S}_a$) based on evidence collected from a single ROI k, if the response feature of the ROI k for the subject ($\dot{S}_a$) was classified into class 1 ($sRI^k \geq 0.5$), the subject was classified as a relapser (class 1). Conversely, if $sRI^k < 0.5$, the subject was classified as an abstainer. For the classification of a subject ($\dot{S}_a$) based on a synthesis of multiple lines of evidence collected from multiple ROIs, in this study, we only focused the investigation on the cases where the classification results of all the observed ROIs in the same subject were in agreement. Then, the classification result of the subject ($\dot{S}_a$) was specified as being the same as one of the classification results of the subject for these ROIs.

For the decision value ($sDec^k$) and probability ($sProb^k$), the methods of their calculation and classification for the response feature of the ROIs k for a subject were applied similarly. The investigation was implemented at the last classification (M = 450) on the 40 alcoholic patients.

Evaluation

In this study, the evaluation and comparison of the classifications on the same dataset were applied in a manner similar to that used in the first and second studies. For the comparison of

two different classifications on the datasets with different sample size, t-test analysis for difference between two proportions in the case of unpaired samples was applied (Zikmund et al., 2013).

Results

1. Classification of the patterns for individual ROIs estimated from the cross-validations

The classification result of pattern for each ROI was averaged over all the test sets created from the 10-repeated 10-fold cross-validation procedure on the 40 alcoholic patients (N = nSF x nB x nF = 4 x 12 x 100; N: number of classified patterns for each ROI k; nSF: number of subjects for each fold; nB: number of blocks for each subjects; nF: number of tested folds). The indices of performance were the average values over all the sets reported (in %).

The results showed that the difference of performance for the right ACC between the two estimates of the decision value and probability was a little significant ($\chi^2_M = 4.5$; p=0.0339) while for the right VS and insula the difference was not significant ($\chi^2_M = 1.32$ and 1.33; p=0.2515 and 0.2482; Table 4.1). Especially, the performance of the right VS, right ACC and insula using SVMs was very significantly higher than that using the simple classification rule based on the difference of the Euclidean distance from the mean vectors of the two classes in the original fMRI space (76.2%, 68.8% and 71.3% vs. 47.9, 48.1 and 50% accuracy respectively; $\chi^2_M = 917.8$, 675.5 and 563.4; p<0.00001; Table 4.1).

Table 4.1. **The classification performance of pattern for individual ROIs**

ROI N = 4800	Classification performance (%) using the distance difference in the original fMRI data				Classification performance (%) using SVMs							
					Decision value				Probability estimate			
	Accuracy	*Sensitivity*	*Specificity*	p_{corr}	*Accuracy*	*Sensitivity*	*Specificity*	p_{corr}	*Accuracy*	*Sensitivity*	*Specificity*	p_{corr}
Right VS	47.9	47.7	48.1	0.9997	76.2	72.9	79.5	<0.0001	75.9	73	78.9	<0.0001
Right ACC	48.1	47.5	48.5	0.9955	68.8	68	69.5	<0.0001	68.2	67.9	68.5	<0.0001
Insula	50	50	50	0.5058	71.3	74	68.5	<0.0001	71.5	74.4	68.6	<0.0001

2. Validation of the ranking

2.1 Validity and reliability of the estimated ranking index

2.1.1 The correlation between the ranking index and the decision value and probability

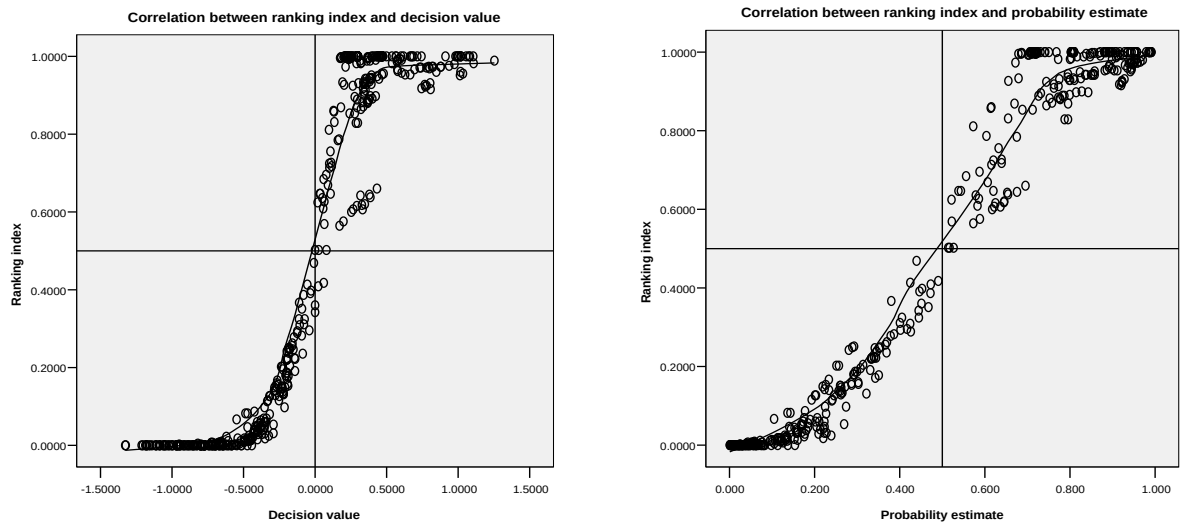
The results showed that all the three estimates correlated strongly and positively (p<0.001; N = nS x nB = 40 x 12; N: number of analyzed patterns for each ROI k; nS: number of subjects; nB: number of blocks for a subject; Table 4.2). In particular, the correlation between the ranking index and probability estimate was very strong and positive (0.978, 0.975 and 0.99 for right VS, right ACC and insula respectively; p<0.001) while the correlation between the ranking index and the decision value was less powerful (0.906, 0.882 and 0.906 respectively; p<0.001; Table 4.2).

Table 4.2. **Correlation between the ranking index & the decision value and probability**

Correlation N = 480	Right VS		Right ACC		Insula	
	*R**	*p-value*	*R*	*p-value*	*R*	*p-value*
Ranking index & Decision value	0.904	<0.001	0.882	<0.001	0.906	<0.001
Ranking index & Probability	0.976	<0.001	0.972	<0.001	0.99	<0.001
Decision value & Probability	0.965	<0.001	0.955	<0.001	0.948	<0.001

**R: Pearson correlation coefficient*

Right Ventral Striatum

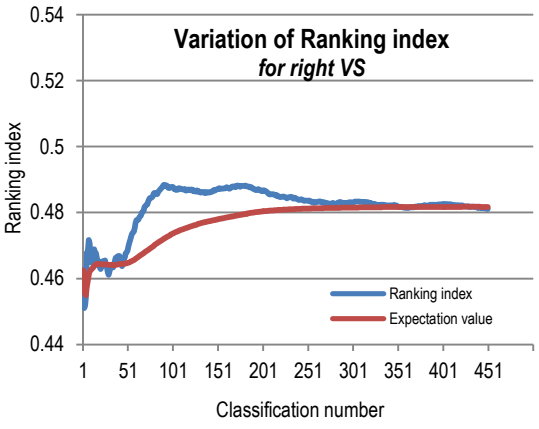


(4.5a) **(4.5b)**

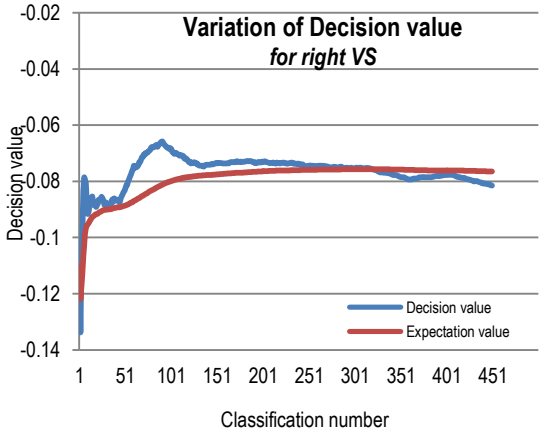
Figure 4.5. **The graphs of the correlation between the ranking index, the decision value and probability** obtained from the analysis on the 480 response patterns of the right VS summarized after 450 classifications. The graph (4.5a) shows that the correlation between the ranking index and the decision value is strong and positive, in particular with regard to the decision values in the range from -0.5 to +0.5. For the decision values outside this range (<-0.5 or >+0.5), most of the ranking indices were approximately equal to 0 and 1 correspondingly. The graph of the correlation between the ranking index and probability estimate (4.5b) indicates that the correlation between them is very strong and positive.

2.1.2 Convergence of the ranking index

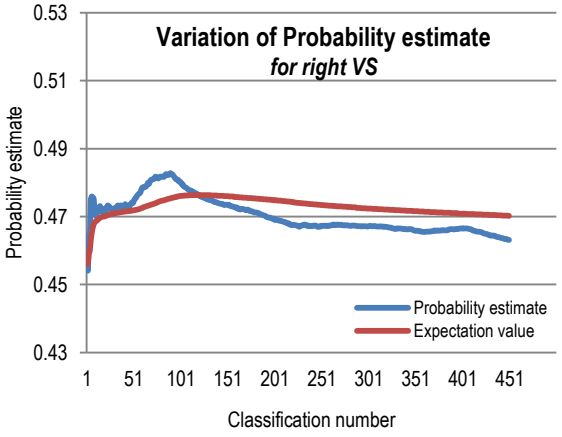
The convergence of the ranking index was investigated through variation of the average ranking index (aRI^j) of the 480-pattern set and variation of the error rate (Er^j) of this index (aRI^j) compared with its expectation value ($E(aRI^M)$) at the M^{th} classification (j: *the* j^{th} *classification*; $j = 1 \rightarrow M; M = 450$). The variation of the average decision value ($aDec^j$) and average probability ($aProb^j$) were considered as well.



(4.6a)



(4.6b)



(4.6c)

Figure 4.6. The graphs of variation of the average ranking index (4.6a), the average decision value (4.6b) and the average probability (4.6c) together with their corresponding expectation values during M classifications (M = 450) for the right VS. The graphs (4.6a, b, c) shows that the ranking index, decision value, probability and their corresponding expectation values fluctuate strongly at the first classifications, and after more than 200 classifications, the amplitude of the fluctuation diminishes gradually. However, while the ranking index and decision value still fluctuate slightly after the 300th classification, their expectation values show convergence. For the right ACC and insula, the ranking index also shows similar convergence. Their graphs are presented in the appendices (Fig. s.1).

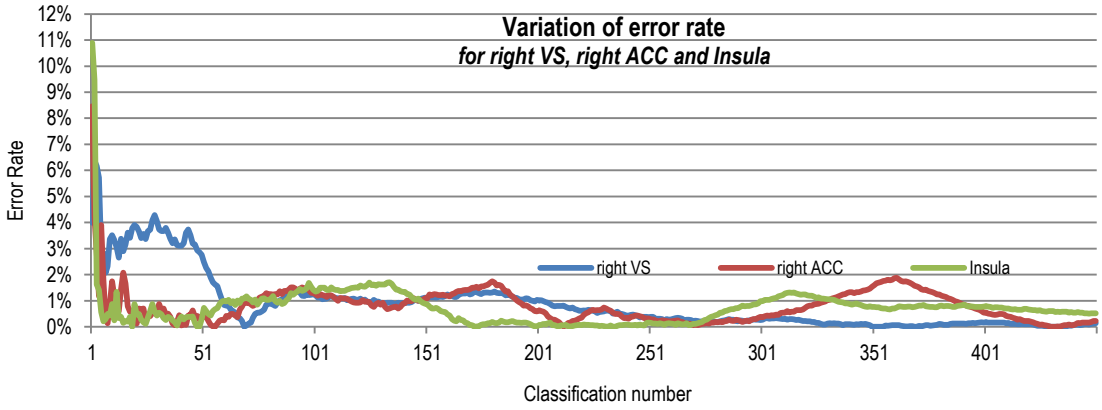


Figure 4.7. The graph of variation of the error rates of the ranking indices (Er^j) for the right VS, right ACC and insula during M classifications compared with their corresponding expectation values at the last classification ($E(aRI^M)$) ($j = 1 \rightarrow M; M = 450$). The graph shows that the error rates of the ranking indices for all the three ROIs fluctuate strongly at the first classifications, and then the amplitude of the fluctuation decreases gradually. After the 400th classification, the error rates fluctuate quite steadily at the low level less than 1%.

2.2 Ranking index & Classification

2.2.1 Classification of the patterns for individual ROIs

For the classifications of the 480-pattern set for individual ROIs based on the ranking index, decision value and probability estimate (Tables 4.3.1 & 4.3.2), the accuracies of classification were said to be significant with 1% level under the Bonferroni correction (p_{corr}<0.01) if their uncorrected p-values were smaller than 0.868 x 10^{-6} (0.01/(Nx*m*); N = 480; *m* = 24; N: number of classified pattern; *m*: number of observations). Equivalently, the accuracies of classification are significant with 1% level if they are greater than 60.8%. On this basis, the results showed that all the classifications for the three investigated ROIs were significant with p_{corr}<0.01; and there was no significant difference between the classification results of the right VS, right ACC and insula in the three methods based on the three indices: the ranking index, the decision value and probability estimate (right VS, $\chi_M^2 = 2.25$, 3.13 and 0.25; right ACC, $\chi_M^2 = 0.25$, 0.8 and 0; insula, $\chi_M^2 = 0.06$, 0, 0; *p*>0.05; N = 40 x 12; Tables 4.3.1 & 4.3.2). Also, these results were not significantly different from those with the evaluation using the cross validation procedure (the right VS, 77.1% vs. 76.2% (z = 0.45; *p*=0.69870); the right ACC, 71.5% vs. 68.8% (z = 1.25; *p*=0.3388); the insula, 70.2% vs. 71.3% accuracy (z = 0.5; *p*=0.665); Tables 4.1 & 4.3.1).

Table 4.3.1. **The classification performance of response patterns (%) for the right VS, right ACC and insula using the ranking index, decision value and probability**

ROI N = 480	Ranking Index (RI)			Averaged decision value $(aDec)$			Averaged probability $(aProb)$		
	Accuracy	*Sensitivity*	*Specificity*	*Accuracy*	*Sensitivity*	*Specificity*	*Accuracy*	*Sensitivity*	*Specificity*
right VS	77.1	78.3	76	77.9	79.7	76.4	78.3	80.6	76.4
right ACC	71.5	71.9	71	71.9	72.7	71.1	72.1	73	71.2
Insula	70.2	68.9	71.8	69.8	69.2	70.4	70	69.4	70.7

There was no significant difference between the classification results based on the ranking index, decision value, probability estimate and their corresponding expectation values (p>0.05; Tables 4.3.1 & 4.3.2).

Table 4.3.2. **The classification performance of response patterns (%) for the right VS, right ACC and insula using the expectation values of the ranking index, decision value and probability**

ROI N = 480	Estimation based on expectation value (E)								
	Ranking Index $(E(RI))$			Averaged decision value $(E(aDec))$			Averaged probability $(E(aProb))$		
	Accuracy	*Sensitivity*	*Specificity*	*Accuracy*	*Sensitivity*	*Specificity*	*Accuracy*	*Sensitivity*	*Specificity*
right VS	77.9	79.7	76.4	78.1	78.7	77.6	77.9	79.1	76.8
right ACC	71.9	72.7	71.1	71.7	72.2	71.1	71.9	71.2	72.7
Insula	69.8	69.2	70.4	70.2	69.3	71.2	70	69.4	70.7

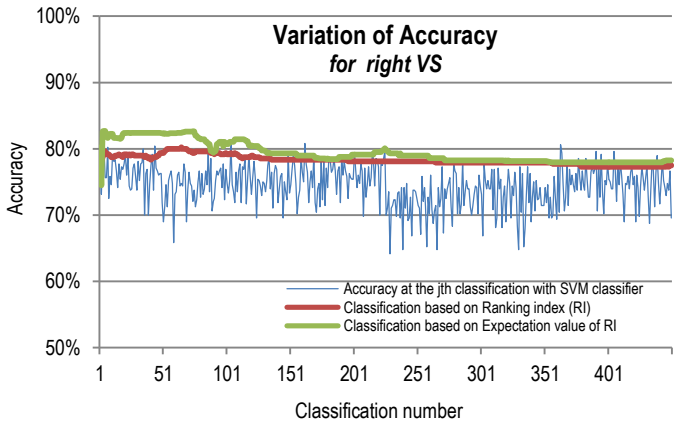


Figure 4.8. The graph of variation of accuracy during M classifications of the response patterns of the right VS (M = 450). The accuracy was calculated based on the ratio of the number of the correctly classified patterns to the total number of classified patterns (N) at the j^{th} classification (N = 480 patterns; j=1→ M; M = 450). The blue line shows the variation of the accuracy during M single classifications using M separate SVM classifiers for each classification. The brown and green lines show the variation of

the accuracies achieved from the classification based on the ranking index (RI) and its expectation value ($E(RI)$) estimated after j classifications. The graph indicates instability of the classification accuracy based on a single classification (blue line) as well as stability of the classification accuracy based on the synthesis of multiple classifications ($1 \rightarrow j$) (brown and green lines) represented by the ranking index as well as its corresponding expectation value.

2.4.2 Classification of subject

2.4.2.1 Based on a single observation result on a single ROI

The results indicated that there was no significant difference between the three methods based on the three indices of the ranking index, decision value and probability (p>0.05; Tables 4.4.1 & 4.4.2). In addition, it seems that the classification results based on the ranking index, decision value and probability were lower than those based on their corresponding expectations (Tables 4.4.1 & 4.4.2). However, the difference was not significant (p>0.05). The detail table of the ranking index (sRI) for the three ROIs for the 40 patients is presented in the appendices (Table s.2).

Table 4.4.1. **The performance of subject classification (%) based on the response feature of a single ROI using the ranking index, decision value and probability**

ROI N = 40	Ranking Index (sRI)				Averaged decision value ($sDec$)				Averaged probability ($sProb$)			
	Accuracy	*Sensitivity*	*Specificity*	*p-value*	*Accuracy*	*Sensitivity*	*Specificity*	*p-value*	*Accuracy*	*Sensitivity*	*Specificity*	*p-value*
right VS	77.5	79	76.2	0.0001	80	80	80	0.00002	80	80	80	0.00002
right ACC	72.5	73.7	71.4	0.0011	72.5	73.7	71.4	0.0011	72.5	73.7	71.4	0.0011
Insula	67.5	66.7	68.4	0.0083	67.5	66.7	68.4	0.0083	67.5	66.7	68.4	0.0083

Table 4.4.2. **The performance of subject classification (%) based on the response feature of a single ROI using the expectation values of the ranking index, decision value and probability**

ROI N = 40	Estimation based on expectation value (E)											
	Ranking Index ($E(sRI)$)				Averaged decision value ($E(sDec)$)				Averaged probability ($E(sProb)$)			
	Accuracy	*Sensitivity*	*Specificity*	*p-value*	*Accuracy*	*Sensitivity*	*Specificity*	*p-value*	*Accuracy*	*Sensitivity*	*Specificity*	*p-value*
right VS	80	80	80	0.00002	80	80	80	0.00002	80	80	80	0.00002
right ACC	72.5	73.7	71.4	0.0011	72.5	73.7	71.4	0.0011	72.5	73.7	71.4	0.0011
Insula	70	70	70	0.0032	70	70	70	0.0032	70	70	70	0.0032

2.4.2.2 Based on multiple observation results on multiple ROIs

The results showed that if we only focused the investigation on the subjects with the agreement classification results on the two ROIs, the classification accuracy was up to over 84% (Table 4.5), and if the classification results on the three ROIs were in agreement, the classification accuracy was 100%. However, in the opposite direction, the number of the excluded subjects increased correspondingly (0 → 14 → 23; Tables 4.5 & s.2).

Table 4.5. **The performance of subject classification based on multiple ROIs**

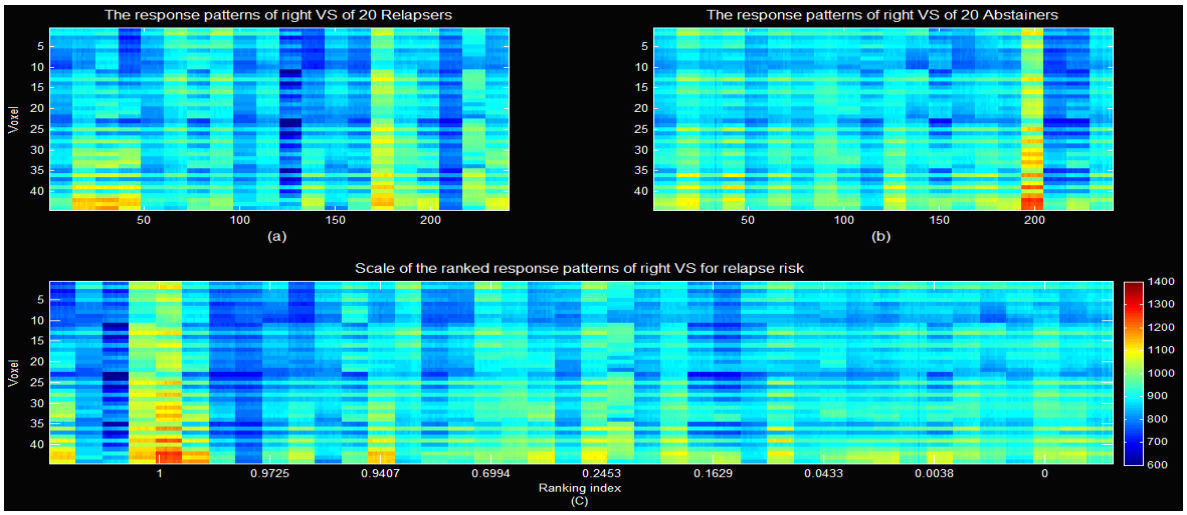
ROI	N = 40		Classification performance (%) on N_1			
	The number of excluded subjects N_2	The number of included subjects N_1	*Accuracy*	*Sensitivity*	*Specificity*	*p-value*
right VS - right ACC	14	26	88.5	91.7	85.7	0.00001
right VS - insula	14	26	84.6	84.6	84.6	0.00027
Insula - right ACC	18	22	86.4	81.8	90.9	0.00006
right VS - right ACC - insula	23	17	100	100	100	0.00000

**N: the number of investigated subjects; N1: the number of subjects with the agreement classification results on the observed ROIs; N2: the number of subjects with the disagreement classification results on the observed ROIs.*

2. *Illustration*

3.1 The ranking of the 480 response patterns of individual ROIs

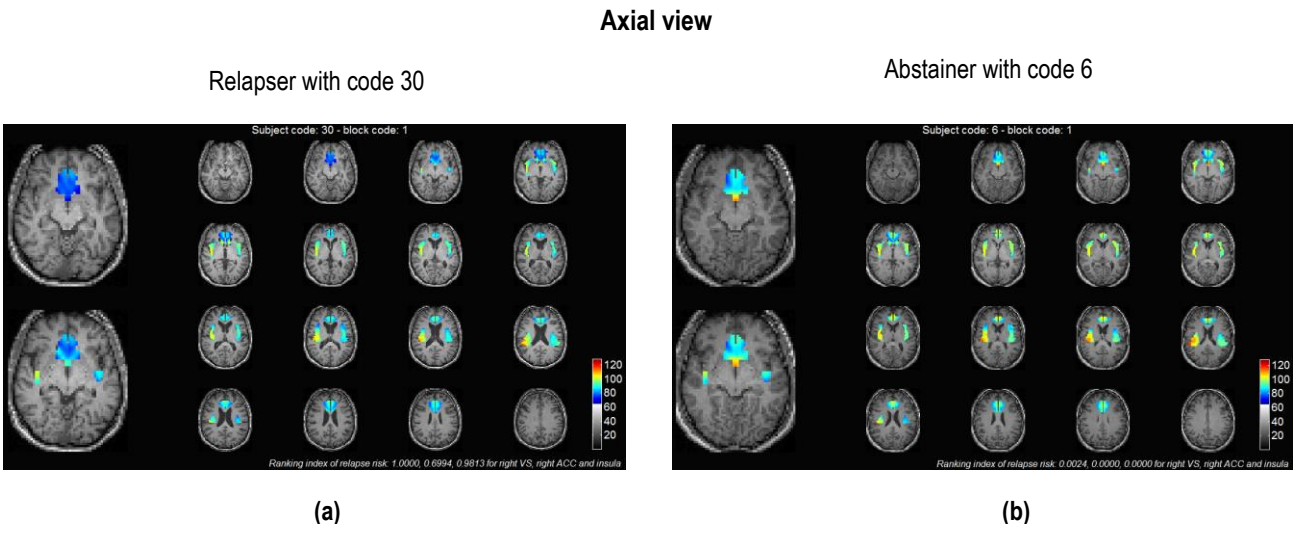
Figure 4.9. The ranking of the 480 response patterns of the right VS to alcohol cues collected from the 20 relapsers (a) and 20 abstainers (b) arranged in the order as shown in Table s.2 (see



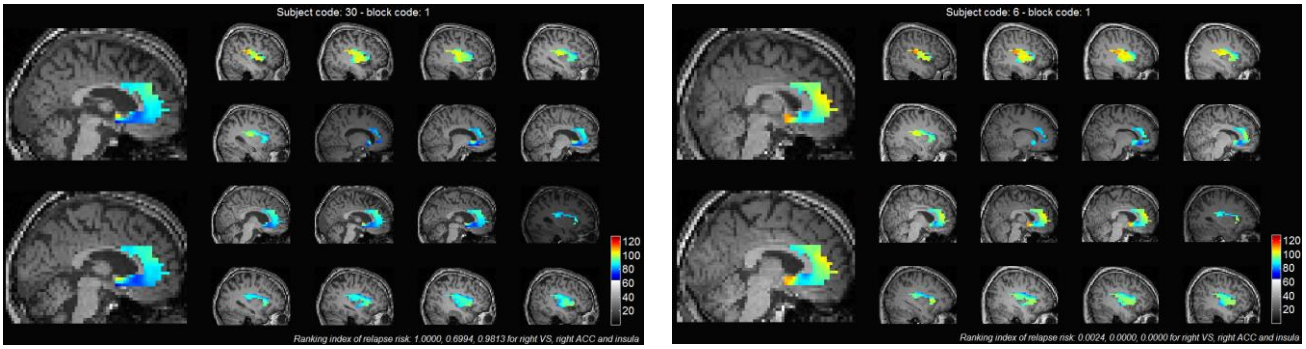
the appendix section). Each column was a response pattern (p_{ki}) including the attributes of all the voxels of the right VS (a_{ki}^{t} ; k: the right VS; $t = 1 \rightarrow v; v$ = 44 (size of the right VS); $i = 1 \rightarrow N$; N = 480) which were arranged in the order of their coordinates (see Fig. 4.3). The response patterns were ranked according to the ranking index of relapse risk for the right VS for subject (sRI^{k}) (c). The figures of the ranking of the right ACC and insula are shown in the appendix section (Figures s.2 & s.3).

3.2 Illustration of the response images of the observed ROIs recorded from 2 alcoholic patients

Figure 4.10. The illustrative response images of the VS, ACC and insula to alcohol-associated cue overlapped on the corresponding structural images acquired from two individual alcoholic patients within the first cue block. The images a, c and e were the images recorded from the relapser with code 30. His ranking indices of relapse risk for the right VS, right ACC and insula were 1, 0.7083 and 0.9807, respectively (Table s.2). The images b, d and f were the images recorded from the abstainer with code 6. His ranking indices for the right VS, right ACC and insula were 0.0024, 0 and 0, respectively (Table s.2). The differences in the activation images recorded on the three ROIs between the two patients can be found in the 14-17th axial slices which were cut through basal ganglia with dimensions of the whole brain image: 53 x 63 x 46 (voxel size, 3 x 3 x 3 mm^3). In the sagittal view, the difference can be observed in the 25-27th sagittal slices (c, d). Further, the differences can be shown more clearly in the 2D and 3D space (e, f). All of the 15th axial and 28th sagittal slices collected from the 40 patients are also shown in the appendix section (Figures s.4 & s.5).

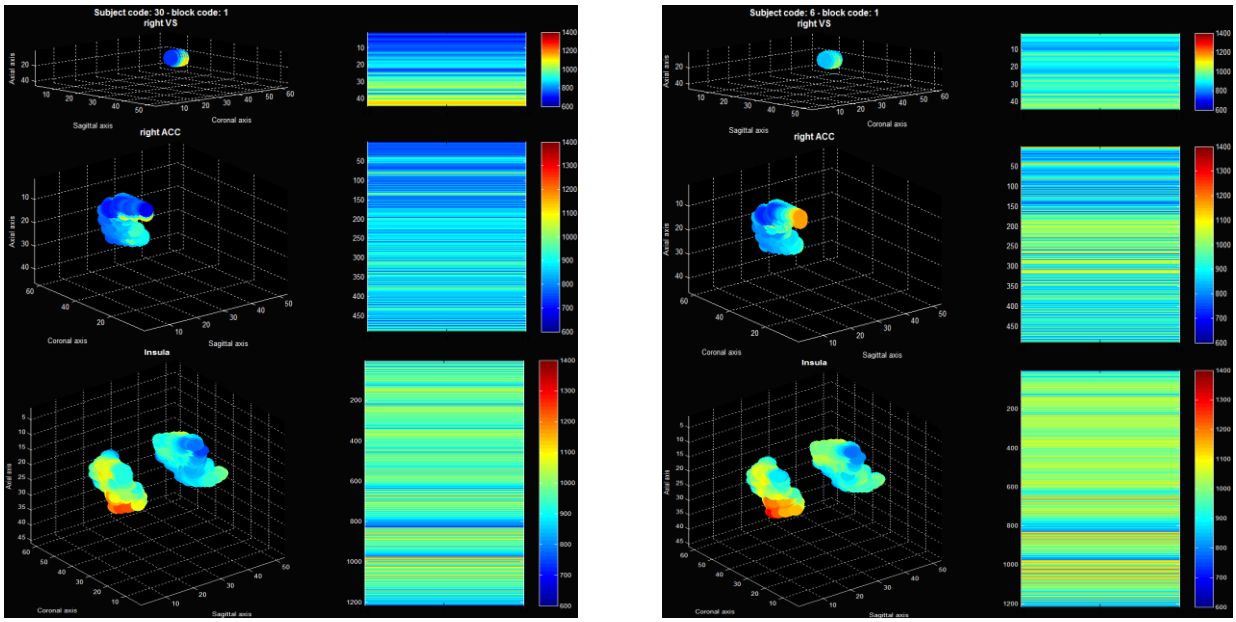


Sagittal view



(c) **(d)**

Cube view



(e) **(f)**

Discussion

Original fMRI data space and SVMs

The classification performance of right VS, right ACC and insula using the method based on the difference of the Euclidean distances from the mean vectors of the two classes in the original fMRI data space were significantly lower than that using SVM classifiers (47.9%, 48.1% and 50% vs. 76%, 69.1% and 70.4% accuracy respectively; p<0.00001; Table 4.1). This indicates

that for fMRI data to get the higher accuracy it is necessary to use classification algorithms. Moreover, in the cases where data are not separable, the classification can be implemented in an appropriately transformed space (Vapnik, 2000; Duda et al. 2001; Wang, 2005; Demirci & Calhoun, 2009). However, such classification methods have involved difficulties in applying diagnostic functional imaging of neuropsychiatric disorders in clinical practice since the basis for the diagnostic imaging is an imaging-based classification method in which the classified images are not transformed in order to keep the anatomical structures of the brain intact (Scott, 2009; Bohland et al., 2012). For this reason, finding the bridges connecting the achievements of information technology solutions with diagnostic imaging has become a practical need. One of the bridges may be found from an understanding of reliability and certainty level of classification decisions for the response patterns obtained from machine inference, because they reflect different levels of response feature between individual patterns, also called discernibility level (Voulgaris & Mirkin, 2008) through which it is possible to reveal the nature of analyzed images forming the basis of diagnostic imaging.

Validation of the ranking algorithm

A single classification and multiple classifications

The reliability of the decision for a pattern (p_{ki}) cannot come from a single classification (Kukar & Kononenko, 2002; Voulgaris & Mirkin, 2008). Let's assume that in a certain unfortunate situation, the support vectors of a SVM classifier identified from training data is inappropriate for classifying the pattern (p_{ki}). In such a situation, the classifier may give a wrong estimate of the decision value for the pattern (p_{ki}) compared with its real nature. This example is similar to an actual case in which mild depression of a particular patient is diagnosed by a surgeon less inaccurately than by a psychiatrist. However, if this patient is examined by hundreds of various specialists, and the diagnosis of his disorder is a synthesis based on all these examinations, the diagnosis would almost achieve high reliability and accuracy reflecting the nature of his disorder (Kasper et al., 2008; Foot et al., 2013).

In the context of SVM classifiers, due to the close dependence of characteristics of SVM classifier on training data, the classifiers formed from different training datasets can be seen as different specialists (Schohn and Cohn, 2000; Vapnik, 2000). Therefore, when the classification

result of a pattern (p_{ki}) is viewed as a measure of its real nature, the measure would be more reliable if it is obtained from multiple classifications (Kukar & Kononenko, 2002; Kimberlin & Winterstein, 2008; Voulgaris & Mirkin, 2008). This hypothesis is supported by the fluctuation of the ranking index (aRI) and its error rate (Er) with large amplitudes in the first classifications (>10%) and with smaller amplitudes when the number of classifications increased (<1% after the 400th; Figures 4.6a & 4.7). This property is also true for output values of the SVM classifier i.e. the decision value ($aDec$) and probability estimate ($aProb$) (Figures 4.6b, c & 4.7). These facts suggest that the measure of reliability of classification for a response pattern does not depend on the classifier but rather mostly on the nature of that pattern or the discernibility level of that pattern when the number of classifications for the pattern increases significantly (Voulgaris & Mirkin, 2008).

Validation of the ranking index

The strong and positive correlation between the ranking index and decision value (the right VS, 0.904; the right ACC, 0.882 and the insula, 0.906; p<0.001; Table 4.2) shows that the ranking index also possesses the characteristic of the decision value, namely that the larger the absolute value, the higher the confidence of classification (Vlachos, 2004; Aydin & Guvenir, 2006). The result is in line with the study of Voulgaris & Magoulas (2008) indicating a positive correlation between the degree of certainty of a classification and the accuracy rate of the classifications. In the context of the study, the former is represented by the absolute decision value, and the latter is represented by the ranking index. Therefore, it can be said that the reliability level of classification for a pattern (p_i) into class 1 (relapser class) is closely related to the ranking index which can be seen as an indicator for the degree of separation of that pattern between the two classes. This interpretation is reinforced by the very strong and positive correlation between the ranking index and probability for the classification into class 1 (the right VS, 0.976; the right ACC, 0.972 and the insula, 0.99; p<0.001; Table 4.2). Another piece of evidence to support the existence of this relationship is that the classification results of pattern based on the ranking index (1) were significantly high (the right VS, 77.1%; the right ACC, 71.5%; the insula, 70.2% accuracy; p<0.0001; Table 4.3.1), and there was no significant difference compared to the results obtained from the 10-repeated 10-fold cross-validation test (2) (the right VS, 77.1% vs. 76.2%;

the right ACC, 71.5% vs. 68.8% and the insula, 70.2% vs. 71.3% accuracy; p=0.698, 0.33 and 0.665 respectively; Tables 4.1 & 4.3.1).

For the reason that the classification results with methods (1) and (2) showed no significant difference in any of the investigated ROIs, the interpretation can be derived from the similarity of the reliability level of estimation between the two approaches. The classification performance obtained from the 10-repeated 10-fold cross-validation test can be seen as a reliable estimate (Payam et al., 2009) while with the design described earlier, the classification performance based on the ranking index at the M^{th} classification (M = 450) can be seen as a variant of the special cross-validation procedure with 'leave one out' which can give an unreliable estimate with large variance (Payam et al., 2009). However, in the latter, the classification decision for a pattern (p_i) did not come from a single classification but rather came from a synthesis of various decisions which gave a more reliable classification decision for that pattern. This interpretation is supported by stability and consistency of the ranking index's measurement characterized by convergence of the ranking index (Kimberlin & Winterstein, 2008) through the gradually stable variation of the ranking index as well as its error rate when the number of classifications increased (Figures 4.6 & 4.7). In addition, another piece of evidence to reinforce the interpretation can be found from the observation of the variation of accuracy during the M classifications that showed the accuracy of the method based on a synthesis of multiple classifications represented by the ranking index to be more stable and consistent than that based on a decision of a single classification (Fig. 4.8).

Taken together, the approach of the study appears to be just a translation from the performance of classifier into the performance of individual response patterns represented by their ranking index (or discernibility level). This is evidenced by the validity of the ranking index in classification as well as the similarity of the performance between the two classification methods (1) and (2). Such a validity and similarity also implies that the higher the performance of a classifier, the higher the validity and accuracy of applying the ranking index produced by the classifier in classification and vice versa. Usefully, this translation provides us an opportunity to obtain the classification decisions accompanied with the corresponding estimates of discernibility level of response patterns which can be used as an indication to identify the difference of response images between the two investigated classes more easily.

Ranking index and methods of estimation

The classification results obtained with the methods based on the ranking index, decision value and probability showed no significant difference (p>0.05; Tables 4.3.1 & 4.4.1). Together with the strong and positive correlation between them, this evidence suggests that not only the classified index into a class but also outputs of SVM classifier such as the decision value or probability estimates can be used as the ranking index as well. Furthermore, they appear to be better when the estimation was based on their corresponding expectation values (e.g. the right VS, 77.5% vs. 80%; the insula, 67.5% vs. 70% accuracy; Tables 4.4.1 & 4.4.2). However, this outperformance needs more evidence because in the study, it did not show a significant difference (p>0.05; Tables 4.3.1 & 4.3.2 and 4.4.1 & 4.4.2), and it seems that these methods can be used interchangeably. Despite this, intuitively the estimate based on the expectation values of the classified index into a class and probability estimate may be an interesting choice because the indices can give us a suggestion of the significance of the discernibility level of the pattern in a more standardized manner while the estimate based on the decision value may produce some unexpected outliers since they are uncalibrated values (Platt, 2000; Vapnik, 2000).

Classification based on multiple ROIs

As the results obtained in the second study, the investigation on the subjects with the agreement classification results on the observed ROIs showed that the accuracy of subject classification increased significantly in the same direction with the number of observed ROIs. For two ROIs, the accuracy increased from 77.5% (right VS), 72.5% (right ACC) to 88.5% (right VS-right ACC), and achieved 100% for three ROIs (right VS-right ACC-insula) (Table 4.5). The evidence suggests that the classification decision for a subject (S_a) may be more reliable and certain when the decision comes from the inference based on multiple evidences collected from several relevant brain regions rather than based on a single piece of evidence from a single region. Such an inference is consistent with the previous fMRI studies of alcohol dependence. The study result of Heinz et al. (2004) indicates the connectivity between the VS and ACC through the inverse correlation between the availability of dopamine receptors in the VS and the strength of activation of the ACC under the stimulation of alcohol-associated cues in alcoholics. The existence of this connectivity is reinforced by the evidence of previous studies of alcohol dependence showing that a significant activation was recorded in both of these regions when

elicited by alcohol-associated stimuli (Grüsser et al., 2004; Myrick et al., 2004; Tapert et al. 2004; Sinha & Li, 2007; Beck et al., 2009; Heinz et al., 2009). Thus, the functional connectivity between them may be the source to generate the selectively different response patterns of the VS and ACC for relapsers and abstainers recorded in the majority of the patients (26/40; Table 4.5). Besides these, the increase of classification accuracy when the number of ROIs increases (100% for the three ROIs; Table 4.5) consolidates the validity of the inference method based on the response features of multiple relevant ROIs.

In the other observation, along with the increase of the performance, the number of excluded subjects due to the disagreement classification results on the observed ROIs increases correspondingly from 0% (40/40) for a single ROI to 35% (14/40) for two ROIs (right VS-right ACC) and 57.5% (23/40) for three ROIs (right VS-right ACC-insula) (Table 4.5). Logically, in a negative perspective, the disagreement can deflect the correct inference direction of several ROIs in some analyzed patterns which can cause misclassifications when the number of the incorrect evidences overwhelms (Stefan & Lionel, 2011). Thus, it can be seen as one of the sources leading to a decline in the accuracy or reliability of classification when the classification is based on multiple lines of evidence. However, in a positive aspect, it may suggest that the disagreement can be seen as a sign indicating that patient may be in an instable status of the investigated disorder which needs to be under an additional follow-up supervision. Therefore, the synthesis of evidence from multiple ROIs can help us discern the various reliability and certainty level of relapse risk from fMRIs. With such a consideration, our hope is to control the decision of machine and to improve reliability of classification as well as to give insights of the brain activity hidden behind the response images. Moreover, this way is compatible with the way of decision-making in clinical practice (Kasper et al., 2008).

Feasibility of imaging diagnosis of the approach

The goal of the ranking is to build the reliably ranked scales of the response patterns of relevant brain regions which can be used as data bank of the response images of these ROIs between the two investigated classes to found for diagnostic imaging.

For example, for a particular pattern (p_i), we can match its response image against its corresponding scale to estimate a position on this scale on which we can predict the class as well as degree of separation between the classes. For this, initially, the classification decision of the

pattern (p_i) received from machine inference is accompanied with an estimate of the discernibility level of the pattern, which can help clinicians approach gradually imaging of the investigated disorder. Let's consider the functional images of right VS, right ACC and insula of the two particular cases one of which was a relapser with code 30, and the other was an abstainer with code 6 (Table s.2). The difference of response images of the right VS and ACC was found in the 14 - 17th axial slices and in the 25 - 27th sagittal slices (Fig. 4.10). If we continue to consider the positions of the response patterns of the right VS for the two patients on the scale, we find that they are located at the 2nd and 34th position with the ranking index of relapse risk = 1 for the subject with code 30 and 0.0024 for the subject with code 6 (Table s.2 & Fig. 4.9). With such an approach, we hope that the interaction between clinician and machine inference would be formed. The machine-based classification decisions become easier to verify and learn, which can help us discern the specific response features of relevant brain regions for the two classes as well as actively exclude ambiguous evidence collected from several ROIs in the inference process to improve the classification accuracy. Then we can step by step approach diagnostic imaging and eliminate our complete dependence on the machine.

Furthermore, with this approach, the generalization of the achievements of the classification to a larger dataset would become the accumulation of experience with the gain that makes the data bank of response images of relevant brain regions more plentiful rather than makes a change of available SVM classifier. Then the generalization can be designed to change the number of SVM classifiers formed from quality training datasets with appropriate size rather than to change the characteristics of the available SVM classifier with a single large training dataset. In addition to reducing dependence of classification decision on a single classifier, such an approach can limit an inversion of SVM optimization when the number of samples of training data increases significantly (H. Yu et al., 2003; Shalev et al., 2008) while retaining the advantages of available good classifiers.

Finally, although the study introduced a feasible approach to realizing diagnostic functional imaging of relapse prediction in practice, there are still numerous challenges ahead in turning the feasibility into reality. To reinforce the approach, we conducted the two pilot studies of clinical application which are presented in the next chapter.

CHAPTER V

FEASIBLE APPLICATIONS IN CLINICAL PRACTICE

Application 1

Feasibility of monitoring treatment response in alcohol dependence using functional imaging

Introduction

In neurological disorders, structural damage to relevant brain regions often exists during a period long enough to be recognized with imaging, and imaging changes of the damage can be used to keep track of treatment response for a specified disorder (Scott, 2009). For psychiatric disorders where the pathology is often due to dysfunction of relevant brain regions, the question this study asks is whether the clues for such a monitoring can be found in the activation patterns of these brain regions. Specifically, the aim of this study was to explore the specific characteristics of activation images for each class of either relapser or abstainer which can be focused on as a sign for monitoring treatment response during detoxification in clinical practice.

Materials and Methods

Materials

This study is a continuation of the third study. The materials and data pre-processing in this study were those used in the third study comprising the 20 relapsers and 20 abstainers. Moreover, only the right VS was chosen for this pilot analysis, and the functional ROI of right VS was just its structural ROI.

Methods

In our three previous studies, characteristics of the patterns for classification were recognized completely by machine. In clinical practice, such characteristics would be recognized through imaging, and for a general preliminary recognition, the characteristics of spectrum or activation level of the patterns can be focused on to observe and analyze. Thus, the objective of this study is

to determine the feasibility of such recognition or to determine whether a significant correlation between the activation level of the patterns and classes exists.

Ranking the pattern according to class (difference of class)

As presented in the methodology of the third study, this ranking indicates the relapse risk for each pattern through the ranking index for the relapser class. Its method is based on class classification. In other words, it is focused on dissimilarity of a pattern (p_{ki}) to the investigated classes with the rule that the pattern can be classified into a class (c_a) if the dissimilarity of the pattern to that class (c_a) is the smallest (Fig. 5.1). With such a classification, the characteristics of differences between the patterns in the same class are of no interest. Therefore, there may be patterns (e.g. p_{ki} and p_{kj}; Fig. 5.1) classified into the same class (c_1) i.e. the same sign ($e.g. f_k(p_{ki}) = f_k(p_{kj}) = sign(w^T \emptyset(p_{ki}) + b) = (sign(w^T \emptyset(p_{kj}) + b) = +1$) with the same confidence i.e. the same decision value ($e.g. (|w^T \emptyset(p_{ki}) + b| = |w^T \emptyset(p_{kj}) + b|$) using SVMs but they may differ in the spectrum or activation level (Fig. 5.1).

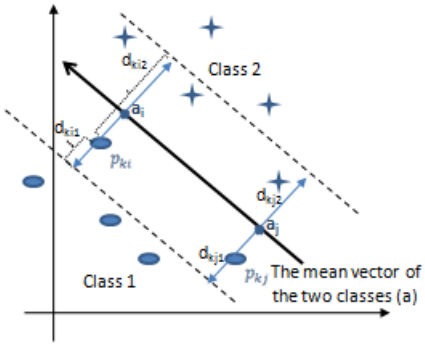


Figure 5.1 **Illustration for class dissimilarity.** p_{ki} and p_{kj}: the response patterns i and j; d_{ki1}, d_{ki2} and d_{kj1}, d_{kj2}: dissimilarities of the patterns p_{ki} and p_{kj} to class 1 and class 2 correspondingly; Given $d_{ki1} = d_{kj1}$ and $d_{ki2} = d_{kj2}$ but the projected coordinates on the mean vector of the two classes (a) of p_{ki} and p_{kj} are different from each other ($a_i \neq a_j$).

Ranking the pattern according to spectrum or activation level (difference of spectrum)

Apart from differences in content, we normally can recognize two different images through a general observation on their spectrum differences. This also means that if the characteristics of spectrum between different classes differ from each other, we can recognize the patterns of the different classes using imaging. For this investigation, we used an average pattern over all the patterns of alcoholic patients comprising the 20 relapsers and 20 abstainers as a standard coordinate axis. Then, the projection position of a pattern (p_{ki}) onto the axis (a_i; Fig. 5.1) can reflect its magnitude on this axis; and this magnitude was used as the ranking index of spectrum or activation level for the pattern (p_{ki}) (Fig. 5.1).

Since the patterns collected from the same subject may possess more similar characteristics than those from the others, in this study, we only analyzed the ranking index of the spectrum for each subject. The index for a subject was the average index over all the corresponding indices of the patterns collected from that subject.

Creating subgroups

After the ranking index of spectrum for each subject had been calculated, all the subjects were sorted according to the index, assigned sequence numbers respectively, and then partitioned into 4 subgroups. Each subgroup comprised 10 different subjects. Subgroup 1 comprised 10 subjects with the 10 lowest indices with the assigned sequence numbers from 1 to 10. Subgroup 2 comprised 10 next subjects with the sequence numbers from 11 to 20. Subgroup 3 consisted of 10 subjects with the sequence numbers from 21 to 30, and subgroup 4 comprised 10 subjects with the numbers from 31 to 40. These 4 subgroups can be considered as the 4 subgroups representing the 4 different levels of the spectrum or activation: the low, low-moderate, strong-moderate and strong level. Then, distribution of the relapsers and abstainers in the subgroups can reveal the correlation between the activation level and classes. For this investigation, all the subjects were labeled as either relapser or abstainer before being distributed in the subgroups. There were two methods of labeling; in one of them the labels of the subjects were specified based on an interview another six months after detoxification as described in the sample collection in the second study (i.e. abstainers: no alcohol consumption; relapsers: any alcohol consumption). In the latter, the subjects were labeled based on the ranking index of relapse risk (the ranking index of class) determined in the third study.

Evaluation

The statistical significance of a proportion between relapsers and abstainers in a subgroup was analyzed based on rejecting the null hypothesis. The null hypothesis assumes that there is no difference between the number of relapsers and abstainers partitioned into the analyzed subgroup. Since the sample size of each subgroup was small ($n<30$), the p-value to reject this hypothesis was estimated based on the binomial distribution as described in the methodology of the first study (Diez et al., 2013). However, here, there was a modification of variables where N trials was the total number of subjects partitioned into the analyzed subgroup, k successful trials

was the number of either relapsers or abstainers in this subgroup according to the analyzed class and the probability of a successful trial p = 0.5. For comparison of the two proportions between the two subgroups, we used t-test analysis for the difference between two proportions in the case of unpaired samples (Zikmund et al., 2013).

Results and Discussion

Correlation between the activation level (spectrum) and classes (relapser and abstainer)

The subjects were labeled based on an interview

The results showed that in subgroup 1 with the low activation level (blue), relapsers prevailed over abstainers (R/n_1=8/10; *p*=0.0107; Table 5.1) while subgroups 2 and 3 with the moderate level (green) mainly comprised abstainers (A/n_{2+3}=14/20; *p*=0.0207; Table 5.1). The ratio of relapsers to abstainers in subgroup 4 with the strong activation level was equal (R/A=5/5; *p*=0.1719; Table 5.1). This appears to indicate that if the activation of the VS is recorded for a subject at the low level, the probability of identifying the subject as a relapser is much higher than that of identifying him as an abstainer. By contrast, the probability of identifying him as a relapser would be lower if this activation is recorded at the stronger level. In other words, the significant difference of distribution of relapsers and abstainers between subgroup 1 and subgroups 2+3 (*p*=0.011) appears to imply that there exists a difference of spectrum or activation level between relapsers and abstainers. The inference is supported by the results of Beck et al. (2012) indicating that the activation of VS in abstainers was stronger than that in relapsers.

Table 5.1. **Correlation between the activation level (spectrum) and classes (relapser and abstainer)**

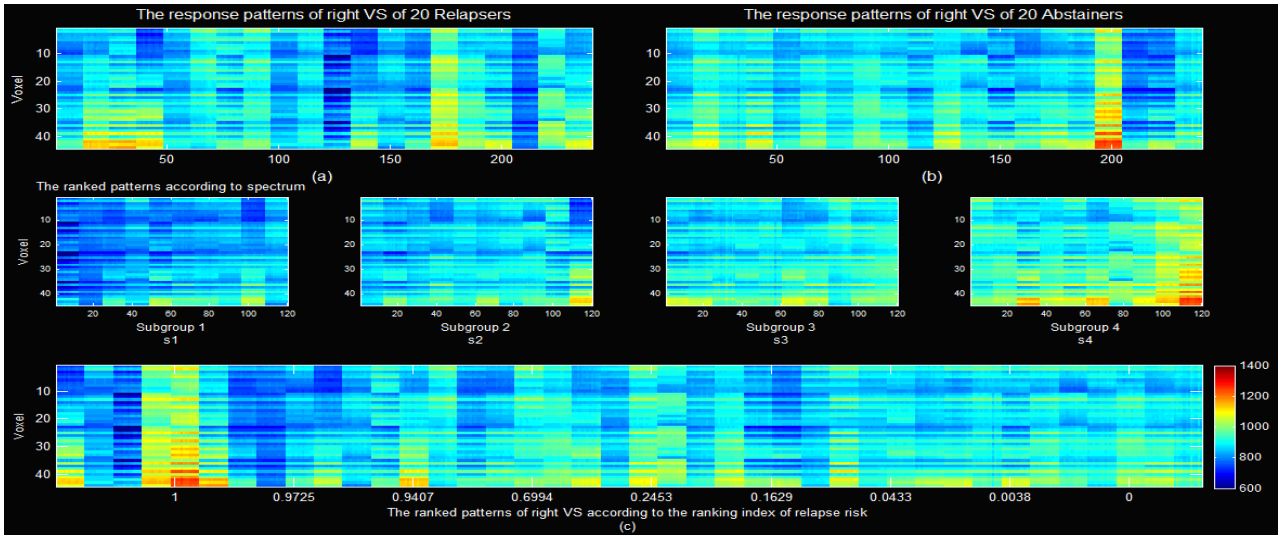
Right VS N = 40	**Activation level (spectrum)**					
	Low (1→10)		**Moderate (11→30)**		**Strong (31→40)**	
	Subgroup 1 n_1 = 10 (1→10)		Subgroup 2 + 3 n_{2+3} = 10 + 10 (11→20 + 21→30)		Subgroup 4 n_4 = 10 (31→40)	
	R/A	*p-value*	*R/A*	*p-value*	*R/A*	*p-value*
Subjects labeled with an interview	8/2	0.0107	3/7 + 3/7	0.0207	6/4	0.1719
Subjects labeled with the ranking index of relapse risk	9/1	0.001	3/7 + 3/7	0.0207	5/5	0.377

**R/A: the number of relapsers (R) and abstainer (A) correspondingly*

The subjects were labeled with the ranking index of relapse risk

The results of the distribution of the subjects with the labeling using the ranking index of relapse risk arrived at in the third study showed no significant difference compared to those with their real labels based on interviews (p>0.5; Table 5.1). This indicates the feasibility of classification combining imaging and decision of machine inference through the ranking index. Usefully, the combination can help us approach gradually imaging of the response feature of the ROIs for different classes through which we can learn the typical response patterns of a ROI for the different classes. Hence, this learning process can give us the ability to check the consistency between the machined-based classification decision for a pattern and its response image with the aim of improving the classification confidence and gradually eliminating the complete dependence on the machine inference.

Figure 5.2. The ranking of the 480 response patterns of the right VS to alcohol cues according to spectrum or activation level. These patterns were collected from 20 relapsers (a) and 20 abstainers (b) arranged in the order as shown in the Table s.2 (see the appendix section). Each column was a response pattern (p_i) including all the attributes of voxels of the right VS as those in Fig. 4.9. The patterns were ranked according to the spectrum or activation level for the subjects and partitioned into the 4 subgroups with the levels of spectrum from low to strong (s1, s2, s3 and s4). Moreover, graph (c) shows the ranked patterns according to the ranking index of relapse risk achieved in the third study.



In this work, we introduced the ranking based on dissimilarities of the spectrum or activation level of the patterns for the right VS. The results showed that there appears to exist a significant correlation between the different subgroups of spectrum and classes. This implies that the characteristics of spectrum of the patterns recorded on this structure from the abstainers are distinguishable from those from the relapsers. Furthermore, the results of Schacht et al. (2011) indicate that the right VS activation in alcoholic patients elicited by alcohol cues appears to be stability across time. Taken together, if abstinence from alcohol is considered as a treatment response in alcoholic patients, it is likely that the characteristics of the spectrum difference between the two classes may be a sign which can be used to monitor this response during detoxification. To confirm this, it is necessary to conduct studies with larger samples where the changes of activation in the brain regions of interest would be monitored before, during and after detoxification.

-*-

Application 2

Feasibility of investigating correlation between clinical variables and functional imaging in alcohol dependence

Introduction

In the third study, the approach is to specify the differences of activation images between relapsers and abstainers through the ranking index of relapse risk (ranking index of class). However, the images characterize vascular changes under the impact of given stimulation rather than directly characterizing neutral activation (Mathews, 2001; Logothetis & Pfeuffer, 2004), while cognitive and behavioral changes of the alcoholic patients are the consequence of neural activation. Therefore, whether correlation between the ranking index and clinical variables exists was the question for this investigation.

Materials and Methods

The materials and the ranking indices of relapse risk for the patterns of the right VS, right ACC and insula collected from the 20 relapsers and 20 abstainers were those used in the third study.

The clinical variable for the investigation was the visual analog rating scale of thirst and hunger (VAS-TH). The rating was performed before and after the fMRI measurement. It ranged from 0 ("no craving at all") to 100 ("severe craving"). The index of the VAS-TH was the sum of the two rating indices of thirst and hunger. As a result, it ranges between 0 and 200.

The investigation of the correlation between the ranking index of relapse risk and clinical variable i.e. VAS-TH was conducted on the two following cases: (1) The analyzed sample included all of the 20 relapsers and 20 abstainers; (2) The analyzed sample only included the subjects who have produced the compatible activation patterns recorded on the combined ROIs i.e. all of the activation patterns recorded on the right VS, right ACC and insula within the same block (extracted from the same pattern of the brain) were classified into the same classes (see the second and third studies).

Results and Discussion

Correlation between the VAS-TH and relapse

We calculated Pearson correlations between the VAS-TH and relapse specified with an interview six months later after detoxification. The results showed no significant correlation for the total sample of the 40 alcoholics before and after the fMRI (R = 0.005 and -0.126; *p*=0.977 and 0.437 respectively; Table 5.2). For the selective subsample of the 17 alcoholics who produced the patterns of the right VS, right ACC and insula with the agreement classification results using their ranking indices, the correlation between the VAS-TH after the fMRI and relapse increased significantly (R = -0.596; *p*=0.012; Table 5.2).

***Table 5.2* Correlation between the VAS-TH and relapse**

Correlation	Whole sample N=40				Selective sample n=17			
	before fMRI		*after fMRI*		*before fMRI*		*after fMRI*	
	*R**	*p-value*	*R*	*p-value*	*R*	*p-value*	*R*	*p-value*
VAS-TH & relapse	0.005	0.977	-0.126	0.437	-0.224	0.387	-0.596	0.012

R:Pearson correlation coefficient*

Correlation between the VAS-TH and ranking index of relapse risk

Again, the results showed no significant correlation between the VAS-TH and ranking indices for the three ROIs before the fMRI measurement for either of the two samples (p>0.05; Table 5.3). However, the correlation became stronger after the fMRI in both samples e.g. the insula: R = -0.373 → -0.674; p=0.14 → 0.003 (n = 17; Table 5.3).

Table 5.3. **Correlation between the VAS-TH and ranking index of relapse risk for functional brain regions**

ROI	Whole sample N=40				Selective sample n=17			
	before fMRI		*after fMRI*		*before fMRI*		*after fMRI*	
	R	*p-value*	*R*	*p-value*	*R*	*p-value*	*R*	*p-value*
Right VS	0.117	0.474	-0.374	0.017	-0.196	0.451	-0.543	0.024
Right ACC	-0.198	0.22	-0.347	0.028	-0.325	0.203	-0.596	0.012
Insula	0.033	0.841	-0.201	0.213	-0.373	0.14	-0.674	0.003

The results appear to indicate that there exists a selective activation on the VS, ACC and insula between relapsers and abstainers during the fMRI experiment involved in thirst and hunger. The agent generating such activation may be derived from the cues used in the experiment. The interpretation is supported by evidence of a significant difference in activation in several brain regions involving the mesolimbic system between alcoholics and healthy controls under impact of alcohol-associated cues reported in previous studies (Braus et al., 2001; Wrase et al., 2007; Park et al., 2007; Beck et al. 2009; Heinz et al., 2009; Schacht et al., 2011; Beck et al., 2012). Therefore, it is likely that when activated specially, the system exposes differences in response in the relevant brain regions between the different classes e.g. alcoholics and controls or relapsers and abstainers which have been identified in the studies.

Compatible activation of the relevant brain regions

The understanding of this activation mechanism has continued to be decoded. However, one of the interesting findings in this study is that the moderate correlations between the VAS-TH and ranking indices of the ROIs were observed in the selective subsample (n = 17; e.g. insula, R=-0.674; p=0.003) while the correlations in the total sample (N = 40) proved weaker (e.g. insula, R=-0.201; p=0.213; Table 5.3). Since the calculation of the ranking index was

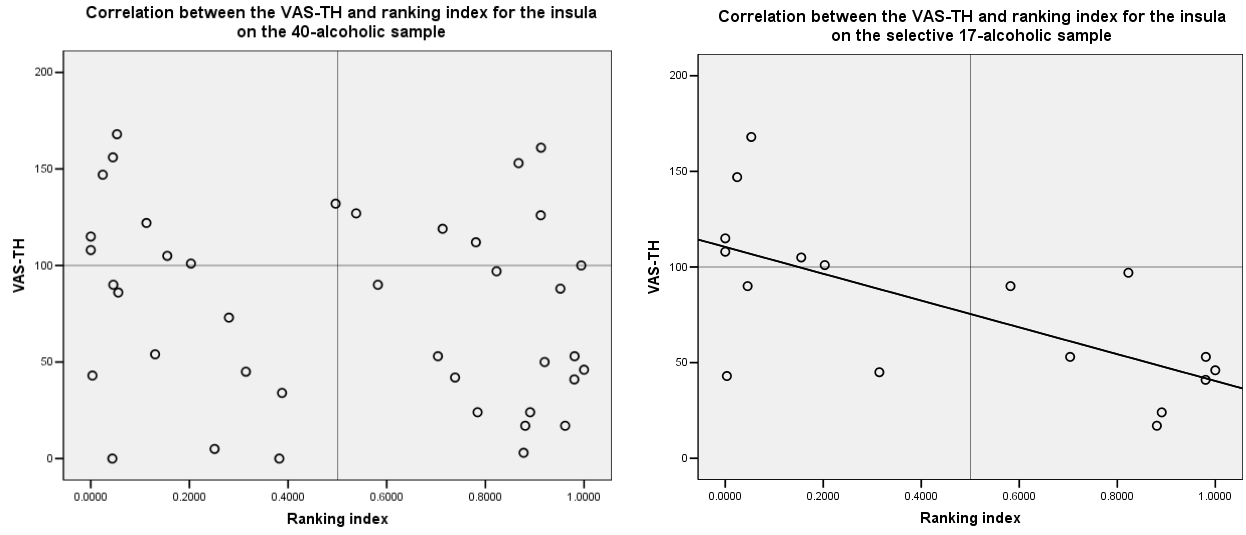


Figure 5.3. **The graphs of the correlation between the VAS-TH and ranking index of relapse risk for the insula.** The rating for the VAS-TH was performed after the fMRI experiment. The graph (5.3a) showed a weak correlation between the two variables on the whole sample (R = -0.201*; p*=0.213; N = 40). The correlation became stronger when estimated only on the selective 17-alcoholic sample (R = -0.674; *p*=0.003; n = 17) (graph 5.3b).

independent of that of VAS-TH, the results appear to imply that there is a bond between the VAS-TH and ranking index which is brought to light more clearly in the selective subsample. Although several lines of evidence have shown that the insula is primarily in charge of homeostatic emotions such as hunger and thirst (Craig, 2003 & 2009), the emotions may be a consequence of the harmonic activation of several relevant brain regions rather than a single region. The basis for the inference is an increase in the strength of the correlations recorded for all of the three ROIs, including the right VS, right ACC and insula in the selective subsample where the activation patterns measured on these three ROIs in the same time were grouped into the same class (Table 5.3). This creates difficulties when specifying the correlation between clinical variables and a response feature of a single brain region. However, the evidence again supports the practicality of the approach to classifying neuropsychiatric disorders as well as their condition using fMRI based on the synthesis of multiple lines of evidence collected from various sources.

CHAPTER VI

GENERAL DISCUSSION AND CONCLUSION

Summary of Discussion

In addition to the understanding of the role of specific brain regions in the diagnosis as well as prediction of relapse risk in alcohol dependence using fMRI presented the earlier chapters, the results support our approach to realizing diagnostic functional imaging in clinical practice. The aspects of location of brain activation, connectivity and classification/prediction achieved from machine inference as well as the decision-making way used by clinicians in the clinical practice of diagnostic imaging have been considered in the approach. Overall, our approach can be summarized with a framework (Fig. 1.8) including the following steps.

(1) Specifying relevant brain regions

The more precise the focus on the brain regions in the aspects of structure and function involved in the pathophysiologic mechanism of the investigated disorder and the cues used to activate them, the higher the classification performance on activation images of the regions using classifiers such as SVMs. Further, the performance may depend on the characteristics of structure and function of the brain regions and their role in the pathophysiologic mechanism rather than on the size of the structures (studies 1 & 2).

(2) Specifying the role of the brain regions of interest in fMRI classification

The classifiers formed from classification algorithms can be used as exploratory instruments to specify the role of the brain regions of interest in fMRI classification in the initial steps (studies 1 & 2).

(3) Translating classification decisions from the classifiers into the discernibility levels between the classes

The translation is to convert classification decisions for the patterns obtained with the classifiers to the discernibility levels (ranking index) showing the separation degree between the two investigated classes. The conversion for each pattern is performed based on the synthesis of the

various classifications for that pattern. The specification of discernibility level for each pattern can be viewed as scoring, and each scored object here is an activation pattern (group of voxels) rather than a single voxel. The data bank of response imaging of brain regions of interest to given stimulation accumulated from the collected subjects can be ranked based on the scores of the patterns. Thanks to the validity of the ranking in classification, we can approach diagnostic functional imaging of the investigated brain regions more easily (study 3).

(4) Classifying an fMRI based on multiple lines of evidence

Like as the way to make a diagnosis of a particular disorder in clinical practice, the inference of a disorder or its condition using fMRI should be based on multiple lines of evidence collected from various sources corresponding to multiple brain regions with different structures and functions, and even from clinical variables in this context. Such an inference can lead to a more reliable decision than those that are only based on a single piece of evidence; and the evidence can be synthesized according to the rule that the more compatible the evidences, the higher the certainty of the classification decision. With such an inference rule, functional connectivity between brain regions and between the brain regions and clinical variables are also considered implicitly. Bayesian inference can be used as a means to realize this rule in the machine-based steps. Since the inference method with Bayes' rule is similar to the decision-making way used by clinicians, the machine inference can be replaced with inference by clinicians in the steps of realizing diagnostic functional imaging in practice (study 3 & application 2).

(5) Approaching activation images of brain regions of interest

For a general observation of imaging, differences of spectrum or activation level of the patterns between different classes are highlighted. The differences can be identified by projecting the patterns onto the same coordinate axis. The distribution of subjects of the classes in the different spectrum groups can reveal rules for this observation (study 3 & application 1).

(6) Considering the compatibility of clinical variables and activation images of relevant brain regions

The logic correlation between clinical variables and activation features of the relevant brain regions is always a major concern in fMRI studies since it would bring a thorough understanding

of the operation of the brain through which we would make an inference of brain disorder more reliable. Although the study result remains limited, it supplies evidence for the feasibility of the application based on the discernibility level (application 2).

With such a framework, the turning-point of the approach to convert the findings of machine-based classification into our understanding of recognition of the analyzed patterns is to translate the machine-based classification decisions for the patterns into indices showing their degree of separation between the two investigated classes. However, the most difficult and key point of the approach to obtain a high achievement in classification is to identify and define as exactly as possible the important brain regions involved in the investigated disorder using fMRI. Such a task is practical for diagnostic imaging because it would help us understand more about the brain and the investigated disorder. Also, it would show the way to identifying and defining the functionally specialized brain regions more exactly. Usefully, the exact definition of the important brain regions impacts positively on the machine-based classification performance and the facilitation for diagnostic imaging approach; and thus a positively reciprocal interaction between us and machine would be formed, which could shorten the road of realizing diagnostic functional imaging in practice.

Limitations

Even if the results of the studies are promising and can be applied to realize diagnostic functional imaging in alcohol dependence, due to the small sample size, the findings of our study may not satisfy clinical diversity in practice. On the other hand, the prediction was only based on a single fMRI measurement after detoxification (1 run), and it was only focused on the differences of the response feature of ROIs between the classes i.e. alcoholics and healthy controls or relapsers and abstainers (between-group) and between different subjects (between-subject). Therefore, the findings might have been more significant and convincing if the response features of the ROIs of each patient were followed up and compared over time before, during and after detoxification along with clinical data of alcohol dependent status (within-subject) on larger samples.

Future works

In this work, we only focused on spatial characteristics of the patterns and ignored the temporal characteristics that may hold useful information for classification. The difference of response to

various cues was not focused on either. Furthermore, in the observation of functional images, we preliminarily examined their differences of spectrum between the classes. Such an observation can be considered as an overview of the images regardless of the important details about content of the images that may hidden the crucial information of differences of the patterns among the classes. Finally, characteristics of the compatibility between the activation images of the relevant brain regions and between the activation images and clinical variables and structural changes of the brain regions of interest e.g. brain atrophy may also keep the valuable information which can help us achieve a logical inference of brain disorders. An inference based on the synthesis from multiple lines of such evidence would significantly improve the certainty of the classification decision. How to unravel these issues will be our major concern in our future works. However, above all, whether and how the approach is fruitful in clinical practice has always occupied a special place in all the questions and solutions that we have been trying to target.

Conclusions

In this dissertation, we proposed an approach to realize diagnostic functional imaging in clinical practice starting from classification algorithm to imaging approach.

In regard to methodology, the approach starts with the formation of functional ROIs from their corresponding anatomical structures and specification of the role of the ROIs in fMRI classification using exploratory instruments formed from the given classification algorithm. The rules of machine-based classification can be transformed into recognition rules of functional imaging in clinical practice. The results demonstrated the feasibility of the approach with a proposed framework.

In regard to alcohol dependence, we found that the VS, ACC and insula are the brain regions playing an important role in relapse prediction using fMRI. The prediction accuracies based on the activation patterns recorded on these brain regions were 63.1%, 58.1% and 71.5% for the VS, ACC and insula respectively (p_{corr}<0.0001). Additionally, the results indicate that there appears to exist a lateralization in response to alcohol cues in the VS and ACC where the activation images of these structures recorded in the right hemisphere were more predictive than those in the left one (e.g. the right VS, 75.9% vs. the left VS, 53%; the right ACC, 68.2% vs. the left ACC, 58.9% accuracy; p<0.0001). For a predictive inference based on multiple observations on

multiple brain regions, a combination of the individual predictions from the right VS, right ACC and insula yielded a better overall prediction (76.9% accuracy; $p<0.0001$).

Although limited in scope, this dissertation helps us to deepen the understanding of relevant brain regions and the method of predictive inference in alcohol dependence using fMRI. Particularly, it provides further evidence that using fMRI images as biomarkers in dealing with neuropsychiatric disorders in clinical practice is feasible.

REFERENCES

Anna Okvist (2009). Synaptic plasticity in drug abuse disorders: studies of the human post-moterm brain. PhD thesis. Departement of clinical neuroscience, Karolinska institutet, Sweden.

Anton R.F., Moak D.H., Latham P. (1995). The Obsessive Compulsive Drinking Scale: A self-rated instrument for the quantification of thoughts about alcohol and drinking behavior. Alcoholism: Clinical and Experimental Research, 19(1):92–99.

Aydin T. & Guvenir H.A. (2006). Modeling interestingness of streaming classification rules as a classification problem. In lecture notes in Computer Science, Springer Berlin / Heidelberg, ISSN 1611-349, 3949, 168-176.

Balakrishnama S. & Ganapathiraju A. (2013). Linear discriminant analysis - a brief tutorial. Accessed March, 2013 at http://www.music.mcgill.ca/~ich/classes/mumt611_07/classifiers/lda_theory.pdf

Barry Horwitz, M-A.Tagamets, Anthony Randal McIntosh (1999). Neural modeling, functional brain imaging, and cognition. Trends in Cognitive Sciences, 3(3):91-98.

Bear F.M., Connors W.B. and Paradisco A.M. (2007). Neuroscience: Exploring the brain. Lippincott Williams & Wilkins (3rd edition).

Beck A., Schlagenhauf F., Wüstenberg T., Hein J., Kienast T., Kahnt T., Schmack K., Hägele C., Knutson B., Heinz A. and Wrase J. (2009). Ventral striatal activation during reward anticipation correlates with impulsivity in alcoholics. Biol Psychiatry, 66:734–742.

Beck A., Grace A.A., and Heinz A. (2011). Reward Processing. In Neuroimaging in addiction (1st edition). Wiley-Blackwell, USA.

Beck A., Wustenberg T., Genauck A., Wrase J., Schlagenhauf F., Smolka N.M., Mann K. and Heinz A. (2012). Effect of brain structure, brain function, and brain connectivity on relapse in alcohol-dependent patients. Arch Gen Psychiatry, 69(8):842-52.

Becker H. (2008). Alcohol dependence, withdrawal and relapse. Alcohol Research, 31(4):343–361.

Berke J.D., Hyman S.E. (2000). Addiction, dopamine, and the molecular mechanisms of memory. Neuron, 25:515–532.

Ben-Hur A., Weston J. (2010). A user's guide to support vector machines. In Data Mining Techniques for the Life Sciences, Methods in Molecular Biology 609. Humana Press, 223-240.

Bland J.M., Altman D.G. (1995). Multiple significance tests: The Bonferroni method. BMJ, 310(6973):170.

Bohland W.J., Saperstein S., Pereira F., Rapin J. and Grady L. (2012). Network, anatomical, and non-imaging measures for the prediction of ADHD diagnosis in individual subjects. Frontiers in Systems Neuroscience, 6(2012).

Bohn M.J.; Krahn D.D.; and Staehler B.A. (1995). Development and initial validation of a measure of drinking urges in abstinent alcoholics. Alcoholism: Clinical and Experimental Research, 19(3):600–606.

Boser B.E., Guyon I.M., Vapnik V.N. (1992). A training algorithm for optimal margin classifiers. D. Proc. Fifth Ann. Workshop on Computational Learning Theory, ACM, 144–152.

Bragulat, V., Dzemidzic, M., Talavage, T. et al. (2008) Alcohol sensitizes cerebral responses to the odors of alcoholic drinks: an fMRI study. Alcohol Clin Exp Res, 32(7):1124–34.

Braus D.F., Wrase J., Grusser S., Hermann D., Ruf M., Flor H., Mann K., Heinz A. (2001). Alcohol-associated stimuli activate the ventral striatum in abstinent alcoholics. Journal of Neural Transmission, 108:887–894.

Büchel C. and Friston K.J. (1997). Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. Cerebral Cortex Dec, 7:768–778.

Bühler M. and Mann K. (2011). Alcohol and the human brain: a systematic review of different neuroimaging methods. Alcohol Clin Exp Res, 35(10):1771–1793.

Burge John (2007). Learning Bayesian networks from hierarchically related data with a neuroimaging application. PhD thesis. The university of New Mexico. Albuquerque, New Mexico.

Castro Norma (2012). The role of the insular cortex in chemosensory responses to ethanol. Accessed on Sep. 2012 at http://sdsu-dspace.calstate.edu/bitstream/handle/10211.10/2414/Castro_Norma.pdf?sequence=1.

Christopher Pickering (2006). Factors affecting alcohol self-administration: learning, environmental, genetic influences on behaviour. PhD thesis. Departement of clinical neuroscience, Karolinska institutet, Sweden.

Clapp P., Bhave V.S. and Hoffman L.P. (2008). How adaptation of the brain to alcohol leads to dependence. Alcohol Res Health, 31(4):310–339.

Clare Stuart (1997). Functional MRI : Methods and applications. PhD thesis. University of Nottingham, UK at http://users.fmrib.ox.ac.uk/~stuart/thesis/fmri.pdf

Cortes C. and Vapnik V. (1995). Support-Vector Networks. Machine Learning, 20:273-297

Craig A.D. (2003). Interoception: the sense of the physiological condition of the body. Curr Opin Neurobiol, 13:500–5.

Craig A.D. (2009). How do you feel — now? The anterior insula and human awareness. Nature Reviews Neuroscience, 10:59-70.

Dahchour A., & De Witte P. (1998). Effect of repeated ethanol withdrawal on microdialysate glutamate release in the hippocampus. Alcohol Clin Exp Res, 22 (A109):175.

David Badre & Mark D'Esposito (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? Nature Reviews Neuroscience, 10:659-669.

Demirci O., Clark V., Calhoun V. (2008a). A projection pursuit algorithm to classify individuals using fMRI data: application to schizophrenia. NeuroImage, 39(4):1774–1782.

Demirci O., Clark V., Magnotta V., Andreasen N., Lauriello J., Kiehl K., Pearlson G., Calhoun V. (2008b). A review of challenges in the use of fmri for disease classification/characterization and a projection pursuit application from a multi-site fMRI schizophrenia study. Brain Imaging and Behavior, 2:207–226.

De Witte Philippe (2004). Imbalance between neuroexcitatory and neuroinhibitory amino acids causes craving for ethanol. Addictive Behaviors, 29:1325–1339.

Diana M., Pistis M., Carboni S., Gessa G.L., Rossetti Z.L. (1993). Profound decrement of mesolimbic dopaminergic neuronal activity during ethanol withdrawal syndrome in rats: electrophysiological and biochemical evidence. Proc Natl Acad Sci USA, 90(17):7966-7969.

Diez D.M., Barr D.C. and Cetinkaya-Rundel M. (2011). OpenIntro Statistics. Accessed on March, 2013 at http://www.openintro.org/stat/down/OpenIntroStatFirst.pdf.

Di chiara G., Imperato A. (1998). Drugs abused by humans preferentially increase synaptic dopamine concentrations in the mesolimbic system of freely moving rats. Proc Natl Acad Sci USA, 85:5274-5278

Dietterich T.G. (1998). Approximate statistical tests for comparing supervised classification learning algorithms. Neural Computation, 10:1895-1923.

Douglas K.M. (2007). A general method for estimating the classification reliability of complex decisions based on configural combinations of multiple assessment scores. PhD thesis. University of Maryland, USA.

Doyon W.M., York J.L., Diaz L.M., Samson H.H., Czachowski C.L., Gonzales R.A. (2003). Dopamine activity in the nucleus accumbens during consummatory phases of oral ethanol self-administration. Alcohol Clin Exp Res, 27(10):1573-82.

Drobes D.J. (2002). Cue reactivity in alcohol and tobacco dependence. Alcohol Clin Exp Res, 26(12):1928−1929.

Drummond D.C. (2000). What does cue-reactivity have to offer clinical research? Addiction, 95(Suppl 2):S129-44.

Duda R.O., Hart P.E., Stork D.G. (2001). Data description and clustering. In Pattern Classification (2nd ed), p. 537:541. Wiley-Interscience, New York.

Edson A.J. & Gareth J.B. (2006). Study design in fMRI: Basic principles. Brain and Cognition, 60(3):220:232.

Etzel J.A., Gazzola V., Keysers C. (2009). An introduction to anatomical ROI-based fMRI classification analysis. Brain Research, 1282:114 – 125.

Eva H.T., Carrie L.M., Elliot T.B., Matthew D.L., and Andrew J.F. (2010).Gaining while giving: An fMRI study of the rewards of familyassistance among White and Latino youth. Soc Neurosci, 5(5-6):508–518.

Everitt J.B., Robbins W.T. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. Nature Neuroscience, 8 (11):1481–1489.

Farah J.M., Gillihan J.S. (2012). Diagnostic brain imaging in psychiatry: Current Uses and Future Prospects. Virtual Mentor, 14(6):464-471. Accessed on March 2013 at http://virtualmentor.ama-assn.org/2012/06/pdf/stas1-1206.pdf.

Filbey F.M., Schacht J., Myers U.M., Chavez R. and Hutchison K.E. (2009). Marijuana craving in the brain. Proceedings of the National Academy of Sciences, 106 (31):13016-13021.

Filbey F.M., Claus E.D. and Hutchison K.E. (2011). A neuroimaging approach to the study of craving. In Neuroimaging in addiction (1st edition). Wiley-Blackwell, USA.

Fisher (1936). The use of multiple measurements in taxonomic problems. Annals of Eugenics, 7:179–188.

Foot C., Naylor C., Imison C. (2013). The quality of general practice diagnosis and referral. Accessed on April 2013 at http://www.isabelhealthcare.com/pdf/Kings_Fund_Diagnosis_and_Referral_2010.pdf.

Fox P.T. and Raichle M.E. (1986). Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. Proc Natl Acad Sci USA, 83:1140–1144.

Friston J.K. (1994). Functional and effective connectivity in neuroimaging: A synthesis. Human Brain Mapping, 2:56-78.

Friston K.J., Ashburner T.J., Kiebel J.T., Nichols E.T., Penny D.W. (2007). Statistical parametric mapping: the analysis of functional brain images. Elsevier, U.K.

Goldstein R.Z., Volkow N.D. (2011). Dysfunction of the prefrontal cortex in addiction: neuroimaging findings and clinical implications. Nat Rev Neurosci, 12:652-669.

Graeme D.J., Regula S.B., Anthony B.W., Gaby S.P.I., and David F.A. (2008). Functional MRI. Graham A. Webb (ed.), Modern Magnetic Resonance, 1037–1050.

Grüsser S.M., Wrase J., Klein S., Hermann D., Smolka M.N., Ruf M., Weber-Fahr W., Flor H., Mann K., Braus D.F., Heinz A. (2004). Cue-induced activation of the striatum and medial prefrontal cortex is associated with subsequent relapse in abstinent alcoholics. Psychopharmacology (Berl), 175:296–302.

Hall Byron (2012). Bayesian Inference. Accessed on Feb. 2012 at http://cran.r-project.org/web/packages/LaplacesDemon/vignettes/ BayesianInference.pdf.

Heinz A., Dufeu P., Kuhn S., Dettling M., Gräf K., Kürten I., Rommelspacher H., Schmidt L.G. (1996). Psychopathological and behavioral correlates of dopaminergic sensitivity in alcohol-dependent patients. Arch Gen Psychiatry, 53:1123–1128.

Heinz A., Siessmeier T., Wrase J., Hermann D., et al. (2004). Correlation between dopamine d2 receptors in the ventral striatum and central processing of alcohol cues and craving. Am J Psychiatry, 161:1783–1789.

Heinz A., Beck A., Wrase J., Mohr J., Obermayer K., Gallinat J., Puls I. (2009). Neurotransmitter systems in alcohol dependence. Pharmacopsychiatry, 42(suppl.1):S95-S101.

Heinz A., Beck A., Grüsser S.M., Grace A.A., and Wrase J. (2009). Identifying the neural circuitry of alcohol craving and relapse vulnerability. Addict Biol, 14(1):108–118.

Heinz A., Beck A., Mir J., Grüsser M.S., Grace A.A., Wrase J. (2010). Alcohol craving and relapse prediction: Imaging studies. In Advances in the neuroscience of addiction (2nd eds). Boca Raton (FL), CRC Press, 137-155.

Helga Höifödt Lidö (2011). Preclinical investigations of Gly-1 inhibition as a new concept for treatment of alcohol dependence. PhD thesis. Addiction Biology Unit, University of Gothenburg, Sweden.

Hojen-Sorensen, Hansen L.K. and Rasmussen C.E. (1999). Bayesian modelling of fMRI time series. Advances in Neural Information Processing Systems 12, S.A. Solla, T.K. Leen and K.-R. Müller (eds.), pp. 754-760, MIT Press (2000).

Hsu C-W, Chang C-C, and Lin C-J (2010). A practical guide to support vector classication. Accessed April, 2011 at http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf.

H. Yu, J. Yang, J. Han (2003). Classifying large data sets using SVMs with hierarchical clusters. In Proceedings of the 9th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 306-315.

Ito R., Dalley J.W., Robbins T.W., Everitt B.J. (2002): Dopamine release in the dorsal striatum during cocaine seeking behavior under the control of a drug-associated cue. J Neurosci, 22:6247–6253.

Jaber Juntu, Jan Sijbers, Steve De Backer, JenyRajan, MTech, and Dirk Van Dyck (2010). Machine learning study of several classifiers trained with texture analysis features to differentiate benign from malignant soft-tissue tumors in T1-MRI images. J Magn Reson Imaging, 31:680–689.

Jones S., Bonci A. (2005). Synaptic plasticity and drug addiction. Current Opinion in Pharmacology, 5: 20–25.

Kalivas P.W., Churchill L., Klitenick M.A. (1993). GABA and enkephalin projection from the nucleus accumbens and ventral pallidum to the ventral tegmental area. Neuroscience, 57(4):1047-1060.

Kalivas P.W. and Volkow N.D. (2005). The neural basis of addiction: a pathology of motivation and choice. Am J Psychiatry, 162:1403–1413.

Kalivas P.W., Volkow N., Seamans J. (2005). Unmanageable motivation in addiction: a pathology in prefrontal-accumbens glutamate transmission. Neuron, 45:647–650.

Kasper D. L., Braunwald E., Fauci A.S., Hauser S.L., Longo D.L., Jameson J.L., Loscalzo J. (2008). Decision-making in clinical medicine. In Harrison's principles of internal medicine (17th ed). New York, McGraw-Hill. 84-103.

Kelley A.E. (1999). Functional specificity of ventral striatal compartments in appetitive behaviours. Ann N Y Accad Sc, 877:71-90.

Kimberlin C.L., Winterstein A.G. (2008). Validity and reliability of measurement instruments used in research. Am J Health Syst Pharm, 65(23):2276-2284.

Koob G.F. and Kreek M.J. (2007). Stress, dysregulation of drug reward pathways, and the transition to drug dependence. Am J Psychiatry, 164:1149-1159.

Koob G.F. and Le Moal M. (2005). Plasticity of reward neurocircuitry and the “dark side” of drug addiction. Natural Neuroscience, 8:1442–1444.

Koob G.F. and Kreek M.J. (2007). Stress, dysregulation of drug reward pathways, and the transition to drug dependence. Am J Psychiatry, 164:1149-1159.

Koob G.F. and Le Moal M. (2008). Addiction and the brain antireward system. Annual Review of Psychology, 59:29–53.

Koob G.F. (2009). Neurobiological substrates for the dark side of compulsivity in addiction. Neuropharmacology, 56: 18–31.

Koob G.F. and Volkow N.D. (2010). Neurocircuitry of addiction. Neuropsychopharmacology Reviews, 35(1):217-238.

Kriegeskorte N., Goebel R., Bandettini P. (2006). Information-based functional brain mapping. PNAS, 103:3863–3868.

Kukar M., Kononenko I. (2002). Reliable classifications with machine learning. In: Proceedings of the 13th European conference on machine learning, pp. 219–231.

Lang P.J., Bradley M.M., Cuthbert B.N. (1999). The International Affective Picture System (IAPS). Gainsville, FL: Center for Research in Psychophysiology, University of Florida.

Lee M.H., Smyser C.D., Shimony J.S. (2012). Resting-State fMRI: A review of methods and clinical applications. AJNR Am J Neuroradiol originally published online on August 30, 2012, 10.3174/ajnr.A3263.

Lee K.H., Brown W.H., Egleston P.N., Green R.D., Farrow T.F., Hunter M.D., Parks R.W., Wilkinson I.D., Spence S.A., Woodruff P.W.R. (2006). A functional magnetic resonance imaging study of social cognition in schizophrenia during an acute episode and after recovery. American Journal of Psychiatry, 163(11):1926-1933.

Lindquist A.M (2008). The statistical analysis of fMRI data. Statistical Science, 23(4):439–464.

Logothetis N.K., Pauls J., Augath M., Trinath T., Oeltermann A. (2001). Neurophysiological investigation of the basis of the fMRI signal. Nature, 412:150–157.

Logothetis N.K., and Pfeuffer J. (2004). On the nature of the BOLD-fMRI contrast mechanism. Magn Reson Imag, 22:1517–1531.

Maldjian J.A., Laurienti P.J., Kraft R.A., Burdette J.H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI datasets. Neuroimage, 19:1233–1239.

Makris N., Oscar-Berman M., Jaffin S.K., Hodge S.M., Kennedy D.N., Caviness V.S., Marinkovic K., Breiter H.C., Gasic G.P., Harris G.J. (2008). Decreased volume of the brain reward system in alcoholism. Biol Psychiatry, 64:192-202.

Mack H.A., Harrington L.A. and Frances J.R. (2010). Clinical manual for treatment of alcoholism and addictions. American Psychiatric Publishing, Inc. London, England.

Mathews M.P. (2001). An introduction to functional magnetic resonance imaging of brain. In Functional magnetic resonance imaging: an introduction to methods. Oxford university press, USA. 3-31.

McClernon F.J., Kozink R.V., Lutz A.M., Rose J.E. (2009). 24-h smoking abstinence potentiates fMRI-BOLD activation to smoking cues in cerebral cortex and dorsal striatum. Psychopharmacology, 204:25–35.

McCune B. & Grace J.B. (2002). Distance measures. In Analysis of Ecological Communities. MjM Software Design, Gleneden Beach, Oregon, pp. 45-57.

Mitchell T.M., Hutchinson R., Niculescu R.S., Pereira F., Wang X., Just M., Newman S. (2004). Learning to decode cognitive states from brain images. Machine Learning, 57:145–175.

Mourão-Miranda J., Bokde A.L.W., Born C., Hampel H., Stetter M. (2005). Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional MRI data. Neuroimage, 28:980–95.

Myrick H., Anton R.F., Li X.B., Henderson S., Drobes D., Voronin K., George M.S. (2004). Differential brain activity in alcoholics and social drinkers to alcohol cues: relationship to craving. Neuropsychopharmacology, 29:393–402.

Naqvi N.H., Bechara A. (2009). The hidden island of addiction: the insula. Trends Neurosci, 32:56–67.

Norman K.A., Polyn S.M., Detre G.J., and Haxby J.V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. TRENDS in Cognitive Sciences, 10(9):424-430.

O'Brien C.P., Childress A.R., Ehrman R., Robbins S.J. (1998). Conditioning factors in drug abuse: can they explain compulsion? J Psychopharmacology 12(1):15-22.

Ogawa S., Lee T.M., Kay A.R., Tank D.W. (1990): Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proc Natl Acad Sci USA, 87:9868–9872.

Oscar-Berman Marlene and Ksenija Marinkovic (2003). Alcoholism and the brain: an overview. Alcohol research & health, 27: 125-133.

O'Toole A.J., Jiang F., Abdi H., Penard N., Dunlop J.P., Parent M.A. (2007).Theoretical, statistical, and practical perspectives on pattern-based classification approaches to the analysis of functional neuroimaging data. J Cogn Neurosci, 19:1735–1752.

Park Mi Sook, Sohn Jin-Hun, Suk Ji-A, Kim Sook-Hee, Sohn Sunju, Sparacio Richard (2007). Brain substrates of craving to alcohol cues in subjects with alcohol use disorder. Alcohol & Alcoholism, 42(5):417–422.

Payam Refaeilzadeh, Lei Tang, Huan Liu (2009). Cross-Validation. Encyclopedia of Database Systems, pp. 532-538.

Pereira F., Botvinick M. (2011). Information mapping with pattern classifiers: a comparative study. Neuroimage, 56(2):476–496.

Pereira F., Mitchell T., Botvinick M. (2009). Machine learning classifiers and fMRI: a tutorial overview. Neuroimage, 45:S199–S209.

Platt J. (2000). Probabilistic outputs for support vector machines and comparison to regularized likelihood methods. In A.J. Smola, P.L. Bartlett, B. Scholkopf, and D. Schuurmans, editors, Advances in Large Margin Classifiers, Cambridge, MA, 2000. MIT Press. Accessed March, 2013 at http://citeseer.nj.nec.com/platt99probabilistic.html.

Rätsch Gunnar (2005). A brief introduction into machine learning. Accessed 10 Jan 2011 at http://events/ccc.de.

Rehm J., Mathers C., Popova S., Thavorncharoensap M., Teerawattananon Y., Patra J. (2009). Global burden of disease and injury and economic cost attributable to alcohol use and alcohol use disorders. Lancet, 373(4):2223–33.

Roggo Y., Duponchel L., Huvenne L.P. (2003). Comparison of supervised pattern recognition methods with McNemar's statistical test: Application to qualitative analysis of sugar beet by near-infrared spectroscopy. Analytica Chimica Acta, 477:187–200.

Room R., Babor T., Rehm J. (2005). Alcohol and public health. Lancet, 365(9458):519-530.

Ross S. and Peselow E. (2000). The Neurobiology of addictive disorders. Clinincal Neuropharmacology, 32:269-276.

Rossetti Z.L., Hmaidan Y., Gessa G.L. (1992). Marked inhibition of mesolimbic dopamine release: a common feature of ethanol, morphine, cocaine and amphetamine abstinence in rats. Eur J Pharmacol, 221(2-3):227-34.

Rossetti Z. L., & Carboni S. (1995). Ethanol withdrawal is associated with increased extracellular glutamate in the rat striatum. European Journal of Pharmacology, 283:177–183.

Schacht P.J., Raymond F.A., Patrick K.R., Xingbao Li, Scott H., Hugh M. (2011). Stability of fMRI striatal response to alcohol cues: A hierarchical linear modeling approach. NeuroImage, 56:61–68.

Schohn G. and Cohn, D. (2000). Less is more: Active learning with support vector machines. In Proceeding of the 17th international conference on machine learning, Morgan Kaufmann, San Francisco, CA, pp. 839–846.

Schultz W., Dayan P., Montague P.R. (1997) A neural substrate of prediction and reward. Science, 275:1593–1599

Schultz Wolfram (1998). Predictive reward signal of dopamine neurons. J Neurophysiol, 80:1-27.

Scott W Atlas (2009). Magnetic resonance imaging of the brain and spine, 4th Edition. Lippincott Williams & Wilkins. Philadelphia. ISBN-13: 978-0-7817-6985-3.

Seghier M.L., Zeidman P., Neufeld N.H., Leff A.P. and Price C.J. (2010) Identifying abnormal connectivity in patients using dynamic causal modeling of fMRI responses. Front Syst Neurosci, 4:142.

Shalev-Shwartz S., Srebro N. (2008). SVM optimization: Inverse dependence on training set size. In Proceedings of the 25th international conference on machine learning (2008), pp. 928-935.

Shinkareva S.V., Ombao H.C., Sutton B.P., Mohanty A., Miller G.A. (2006). Classification of functional brain images with a spatio-temporal dissimilarity map. NeuroImage, 33(1):63–71.

Sinha R., Li C.S.R. (2007). Imaging stress- and cue-induced drug and alcohol craving: association with relapse and clinical implications. Drug and Alcohol Review, 26:25–31.

Skinner H.A. and Horn J.L. (1984). Alcohol Dependence Scale (ADS): User's Guide, Toronto: Addiction Research Foundation.

Schölkopf B., Smola A. (2000). Learning with kernels. MIT Press.

Song-yun Xie, Rong Guo, Ning-fei Li, Ge Wang, Hai-tao Zhao (2009). Brain fMRI processing and classification based on combination of PCA and SVM. International Symposium on Neural Networks - ISNN , pp. 3384-3389.

Spanagel R. (2003). The role of the glutamatergic system in alcohol addiction. Fortschr Neurol Psychiatr, 71:S33-35.

Stefan Conrady and Lionel Jouffe (2011). Introduction to Bayesian networks. Accessed on Feb. 2011 at http://conradyscience.com/white_papers/BBN_Introduction_V13.pdf.

Steven E.H., Robert C.M. and Eric J.N. (2006). Neural mechanisms of addiction: the role of reward-related learning and memory. Annu Rev Neurosci, 29:565–98.

Takayanagi Y., Takahashi T., Orikabe L., Mozue Y., Kawasaki Y., Nakamura K., Sato Y., Itokawa M., Yamasue H., Kasai K., Kurachi M., Okazaki Y., Suzuki M. (2011). Classification of first-episode schizophrenia patients and healthy subjects by automated MRI measures of regional brain volume and cortical thickness. PLoS One 6(6), e21047.

Tapert S.F., Brown G.G., Baratta M.V., Brown S.A. (2004). fMRI BOLD response to alcohol stimuli in alcohol dependent young women. Addict Behav, 29:33–50.

T.-F. Wu, C.-J. Lin and R. C. Weng (2004). Probability estimates for multi-class classification by pairwise coupling. Journal of Machine Learning Research, 5:975-1005.

Thomas F.B., John C.H.B., John B.S., Maristela G.M. (2001). The alcohol use disorders identification test, guidelines for use in primary care. World Health Organization. Accessed March 2011 at http://whqlibdoc.who.int/hq/2001/WHO_MSD_MSB_01.6a.pdf.

Tononi G., Sporns O., Edelman G.M. (1994). A measure for brain complexity: relating functional segregation and integration in the nervous system. Proc Natl Acad Sci USA, 91:5033-5037.

Tzourio-Mazoyer N., Landeau B., Papathanassiou D., Crivello F., Etard O., Delcroix N., Mazoyer B., Joliot M. (2002). Automated anatomical labeling of the MNI MRI single-subject brain. Neuroimage, 15:273–289.

Valiant L.G. (1984). A theory of the learnable. Communications of the ACM 27(11):1134-1142.

Van Horn J.D., Poldrack R. A. (2009). Functional MRI at the crossroads. Int J Psychophysiol, 73(1):3-9.

Vapnik V.N. (1998). Statistical learning theory. Wiley, Canada.

Vapnik V.N. (2000). The nature of statistical learning theory. Springer, New York, NY.

Vincent P.C., Gregory K.B., Robert E.A., Piyadassa K., John P.P., Terran D.R.L., Kent A.K., Vince D.C. (2012). Reduced fMRI activity predicts relapse in patients recovering from stimulant dependence. Human Brain Mapping. doi: 10.1002/hbm.22184.

Vlachos A. (2013). Active Learning with Support Vector Machines. Accessed on March, 2013 at http://www.cl.cam.ac.uk/~av308/vlachos_msc_thesis.pdf.

Volkow N.D., Fowler J.S. and Wang G.S. (2003). The addicted human brain: insights from imaging studies. J Clin Invest, 111(10):1444–1451.

Vollstädt-Klein S., Svenja W., Juri R., Mira B., Oliver K., Gabriele E., Derik H. and Karl M. (2010). Initial, habitual and compulsive alcohol use is characterized by a shift of cue processing from ventral to dorsal striatum. Addiction, 105:1741–1749.

Voulgaris Z. & Magoulas G.D. (2008a). Extensions of the k nearest neighbour methods for classification problems. In the 26th IASTED Conference on Artificial Intelligence and Applications, 23-28. Innsbruck, AU.

Voulgaris Z., Mirkin B. (2008b). Optimising a reliability measure for classification. In Proceedings of the 2008 UK Workshop on Computational Intelligence.

Wang Lipo (2005). Support Vector Machines: Theory and applications. Springer-Verlag, Berlin Heidelberg. pp 1–49.

Weishaupt D., Köchli V.D., Marincek B. (2008). How does MRI work? An introduction to the physics and function of magnetic resonance imaging (2nd edition). Springer-Verlag, Berlin Heidelberg.

Weiss F., Lorang M.T., Bloom F.E., Koob G.F. (1993). Oral alcohol self-administration stimulates dopamine release in the rat nucleus accumbens: genetic and motivational determinants. J Pharmacol Exp Ther., 267(1):250-8.

Weiss F., Parsons L.H., Schulteis G., Hyytiä P., Lorang M.T., Bloom F.E., Koob G.F. (1996). Ethanol self-administration restores withdrawal-associated deficiencies in accumbal dopamine and 5-hydroxytryptamine release in dependent rats. J Neurosci, 16:3474–3485.

William White, Ernest Kurtz and Caroline Acker (2011). The combined addiction disease chronologies 2001. Accessed on Feb 2011 at http://www.williamwhitepapers.com/pr/2001Addiction%20as%20Disease%20 Chronology.pdf.

Wise R.A. (1998). Drug-activation of brain reward pathways. Drug Alcohol Dependence. 51:13-22.

Wrase J., Schlagenhauf F., Kienast T., Wüstenberg T., Bermpohl F., Kahnt T., Beck A., Ströhle A., Juckel G., Knutson B., Heinz A. (2007). Dysfunction of reward processing correlates with alcohol craving in detoxified alcoholics. NeuroImage, 35:787–794.

Yang Y., Chefer S., Geng X., Gu H., Chen X., and Stein E.A. (2011). Structural and functional neuroimaging methods: applications to substance abuse and addiction, in Neuroimaging in Addiction (eds B. Adinoff and E. A. Stein), John Wiley & Sons, Ltd, Chichester, UK.

Ye Yang (2010). Pattern analysis of functionnal MRI data. Master thesis. Texas Tech University.

Zikmund G.W., Babin J.B., Carr C.J., Griffin M. (2013) . Bivariate statistical analysis: differences between two variables. Accessed on March, 2013 at http://www.cengage.com/marketing/book_content/1439080674_zikmund/ book/ch22.pdf.

APPENDICES

***Table s.1.*Clinical data of the 40 alcoholic patients**

No.	Code	Real class	Gender	Age	Age at onset	ADS	OCDS	AUQ	VAS-TH before fMRI	VAS-TH after fMRI	Number of abstinent days until relapse
1	27	1	0	38	28	32	21	23	52	73	34
2	28	1	0	48	45	12	27	16	125	50	18
3	29	1	0	56	36	27	28	21	182	126	29
4	**30**	**1**	**0**	**43**	**35**	**26**	**14**	**16**	**42**	**53**	**154**
5	31	1	0	42		18	23	16	29	17	43
6	32	1	0	30		15	9	37	28	24	23
7	33	1	0	49	37	3	13	17	43	53	20
8	34	1	0	35	35	13	13	21	4	41	40
9	35	1	1	49	34	15	18	16	12	46	179
10	36	1	1	33	33	11	17	17	102	161	75
11	42	1	0	64	20	13	14	16	2	5	38
12	43	1	0	47		10	11	16	30	34	25
13	44	1	0	32	32	16	17	23	61	97	141
14	45	1	0	43	43	10	17	25	0	17	90
15	46	1	0	23	23	12	8	24	101	112	42
16	47	1	1	57	47	17	12	16	10	153	31
17	48	1	0	61	41	16	15	32	55	156	79
18	49	1	1	58	55	18	16	16	67	90	51
19	106	1	1	25		8	12	16	8	0	74
20	107	1	1	48	38	15	24	43	57	132	33
21	1	-1	0	26		14	9	22	6	24	
22	2	-1	0	40	37	17	18	16	136	127	
23	3	-1	0	51	48	12	26	16	26	43	
24	4	-1	0	49	39	18	10	16	59	100	
25	5	-1	0	43	28	13	15	16	9	105	
26	**6**	**-1**	**0**	**28**	**27**	**24**	**17**	**30**	**112**	**108**	
27	7	-1	0	46	20	9	12	18	5	88	
28	*8*	-1	0	54	42	10	16	16	54	54	
29	9	-1	1	57	53	10	10	16	25	115	
30	10	-1	1	28	28	25	19	40	30	147	
31	37	-1	0	49	31	22	15	22	102	119	

32	38	-1	0	48	39	6	8	34	97	168	
33	39	-1	0	50		6	12	22	46	86	
34	40	-1	1	54	40	22	22	16	96	122	
35	41	-1	1	40	32	14	9	28	39	101	
36	51	-1	0	30	30	28	19	116	7	45	
37	53	-1	0	59	56	15	7	16	6	3	
38	54	-1	1	62	35	13	19	16	73	42	
39	55	-1	0	41	41	28	26	45	72	90	
40	100	-1	0	33	26	15	14	16	2	0	

**Note:Class*: 1: relapser, -1: abstainer; Gender: 0: male; 1: female;

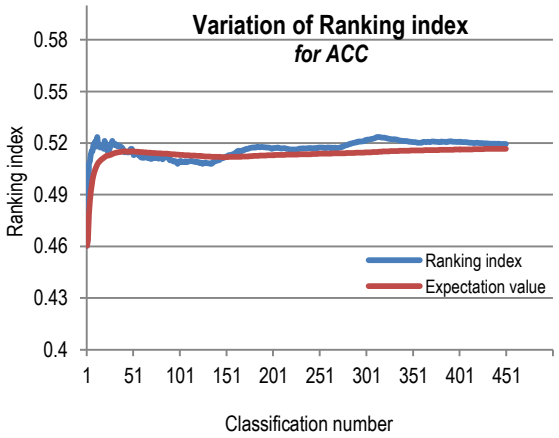
***Table s.2.*The table of the ranking index (*sRI*) of relapse risk for the right VS, right ACC and insula for the 40 alcoholic patients**

No.	Code	Real class	Right VS			Right ACC			Insula		
			sRI	*E(sRI)*	*Predict*	*sRI*	*E(sRI)*	*Predict*	*sRI*	*E(sRI)*	*Predict*
1	27	1	0.9726	0.9915	1	1	1	1	0.28	0.3461	-1
2	28	1	0.9407	0.9538	1	0.3926	0.3914	-1	0.92	0.9207	1
3	29	1	0.9998	1	1	0.0969	0.0636	-1	0.9122	0.9168	1
4	**30**	**1**	**1**	**1**	**1**	**0.7083**	**0.7116**	**1**	**0.9807**	**0.985**	**1**
5	31	1	0.4939	0.533	-1	0.9907	0.9912	1	0.9617	0.9727	1
6	32	1	0.9239	0.9461	1	0.7117	0.688	1	0.8907	0.8631	1
7	33	1	0.9469	0.943	1	0.9994	0.9998	1	0.7037	0.6505	1
8	34	1	0.6994	0.7059	1	0.9841	0.9871	1	0.9802	0.9882	1
9	35	1	1	1	1	0.8802	0.9121	1	1	1	1
10	36	1	0	0	-1	0.2361	0.186	-1	0.9128	0.8993	1
11	42	1	1	1	1	0.2017	0.1868	-1	0.2509	0.2518	-1
12	43	1	0.9683	0.9933	1	0.8341	0.9232	1	0.3876	0.4334	-1
13	44	1	0.948	0.9306	1	0.7157	0.764	1	0.8226	0.8138	1
14	45	1	0.8633	0.9553	1	0.8333	0.8059	1	0.8807	0.8799	1
15	46	1	1	1	1	0.2126	0.1281	-1	0.7807	0.8251	1
16	47	1	0	0.0003	-1	0.0026	0.0008	-1	0.8672	0.8389	1
17	48	1	0.003	0	-1	0.8002	0.7306	1	0.0448	0.062	-1
18	49	1	0.9846	0.999	1	0.9572	0.9604	1	0.582	0.5107	1
19	106	1	0.2331	0.1583	-1	1	1	1	0.3822	0.4366	-1
20	107	1	0.5194	0.5504	1	0.7852	0.7884	1	0.4963	0.3413	-1
21	1	-1	0.0124	0.0043	-1	0.8709	0.8831	1	0.7841	0.8111	1
22	2	-1	0.2454	0.2719	-1	0.955	0.9243	1	0.538	0.4617	1

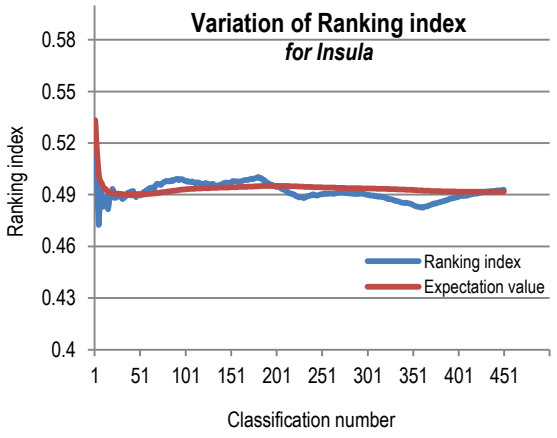
23	3	-1	0.0039	0.0015	-1	0.3846	0.4043	-1	0.0033	0.0017	-1
24	4	-1	0.082	0.0445	-1	0.0041	0.0006	-1	0.9943	0.9964	1
25	5	-1	0.2331	0.2038	-1	0.0646	0.0664	-1	0.155	0.1586	-1
26	**6**	**-1**	**0.0024**	**0.002**	**-1**	**0**	**0**	**-1**	**0**	**0**	**-1**
27	7	-1	0.0433	0.0211	-1	0.2137	0.1772	-1	0.952	0.9406	1
28	*8*	-1	0.6031	0.5114	1	0.007	0.0133	-1	0.1304	0.1516	-1
29	9	-1	0.0202	0.0197	-1	0.2	0.2515	-1	0	0	-1
30	10	-1	0.1193	0.1547	-1	0.2131	0.217	-1	0.0243	0.026	-1
31	37	-1	0.2226	0.1731	-1	0.0165	0.0076	-1	0.7133	0.7636	1
32	38	-1	0.0024	0.0016	-1	0	0	-1	0.0531	0.0966	-1
33	39	-1	0.163	0.1359	-1	0.9994	0.9998	1	0.0557	0.0599	-1
34	40	-1	0.8633	0.8846	1	0.8046	0.8104	1	0.1128	0.0812	-1
35	41	-1	0	0	-1	0.3383	0.3751	-1	0.203	0.1737	-1
36	51	-1	0.0063	0.0074	-1	0	0	-1	0.3144	0.3082	-1
37	53	-1	1	1	1	0	0	-1	0.8774	0.9127	1
38	54	-1	0.9946	0.997	1	0.3996	0.3683	-1	0.7385	0.7391	1
39	55	-1	0.1489	0.1748	-1	0	0	-1	0.0456	0.0201	-1
40	100	-1	0	0	-1	0.9015	0.9517	1	0.0437	0.0301	-1

**Note: sRI*: the ranking index for subject; $E(sRI)$: the expectation value of ranking index for subject; *Predict*: predicting response feature of a ROI for a subject with the rule that if the *sRI* for that ROI is greater or equal to 0.5 ($sRI \geq 0.5$), the response feature of the ROI for that subject was classified into class 1 and by contrast, if $sRI < 0.5$, the response feature of the ROI for that subject was classified into class 2.

Figure s.1. The graphs of variation of the average ranking index (aRI) (blue line) and its corresponding expectation values ($E(aRI)$) (brown line) during M classifications (M = 450) for the right ACC (Fig. s.1a) and insula (Fig. s.1b).



(s.1a)



(s.1b)

Figure s.2. The ranking for the 480 response patterns of the right ACC to alcohol cues collected from the 20 relapsers (a) and 20 abstainers (b) (Tables s.1 & s.2).

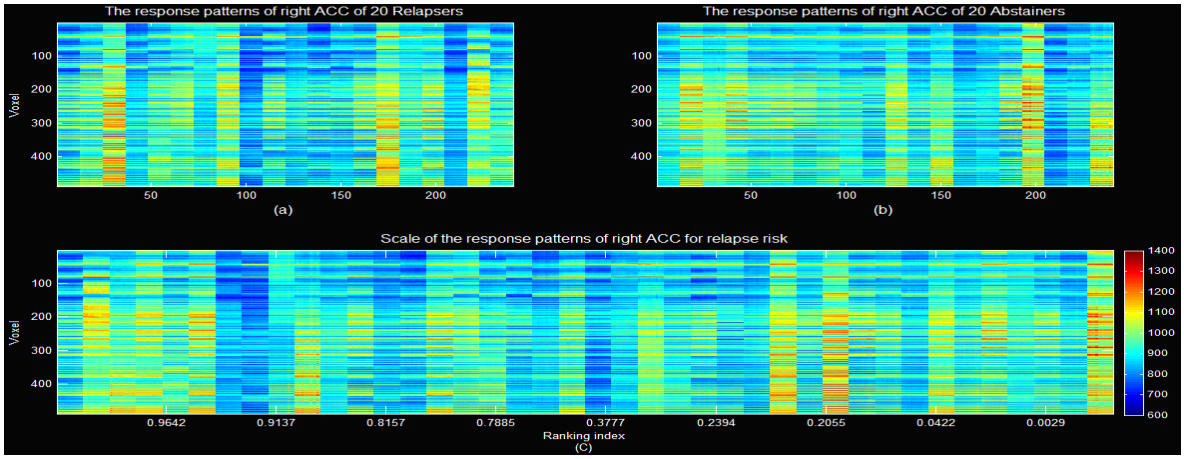


Figure s.3. The ranking for the 480 response patterns of the insula to alcohol cues collected from the 20 relapsers (a) and 20 abstainers (b) (Tables s.1 & s.2).

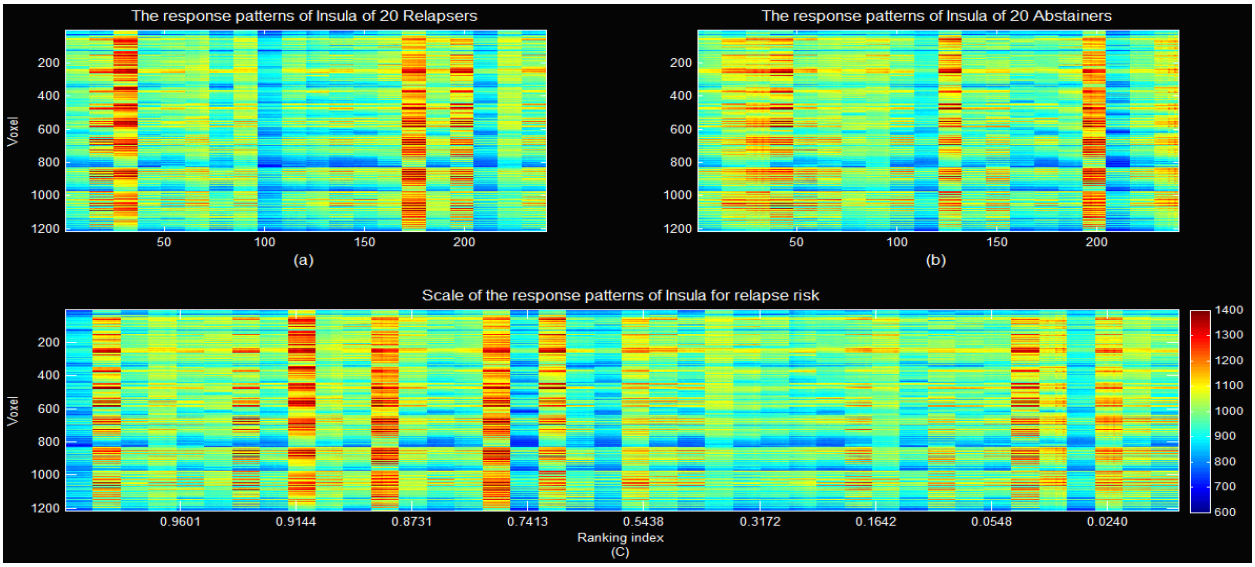
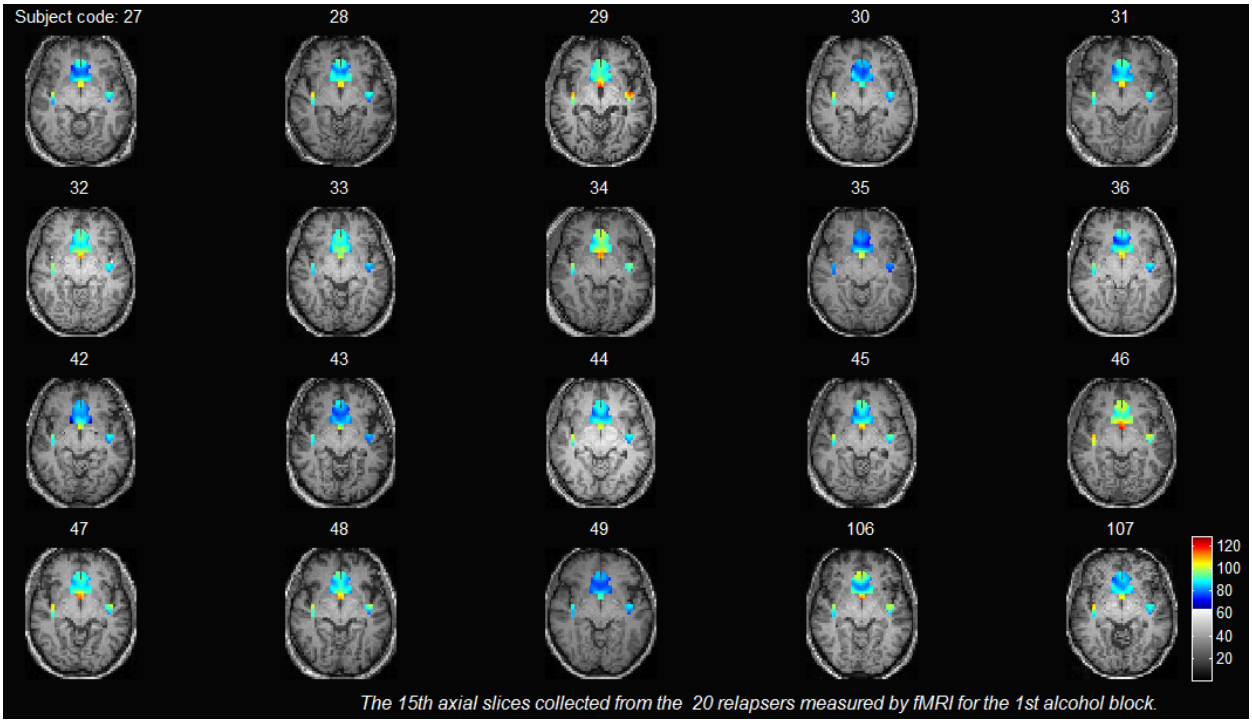
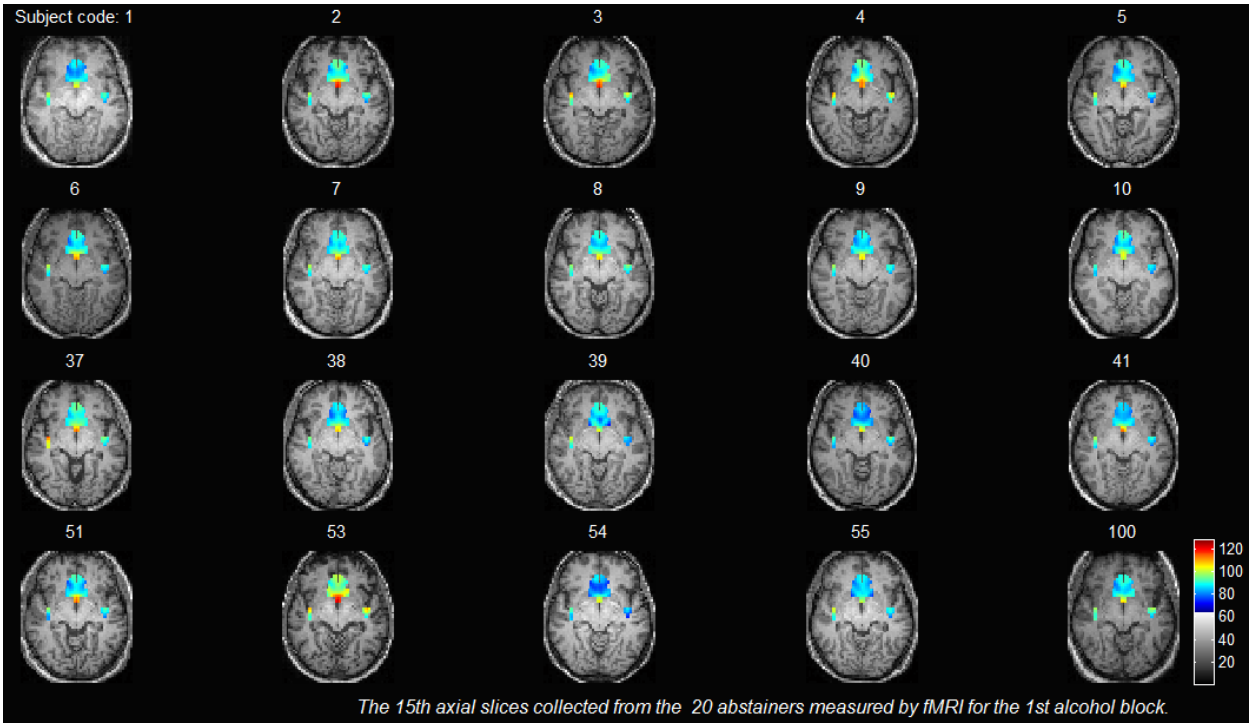


Figure s.4. The response images of the VS, ACC and insula to alcohol-associated cues at the 15th axial slice (a, b) and 28th sagittal slice (c, d) overlapped on the corresponding structural images acquired from the 20 relapsers (a, c) and 20 abstainers (b, d) (Tables s.1 & s.2) within the first cue block. Each slice corresponds to each subject. The code of each subject are attached to each slice.

Axial view

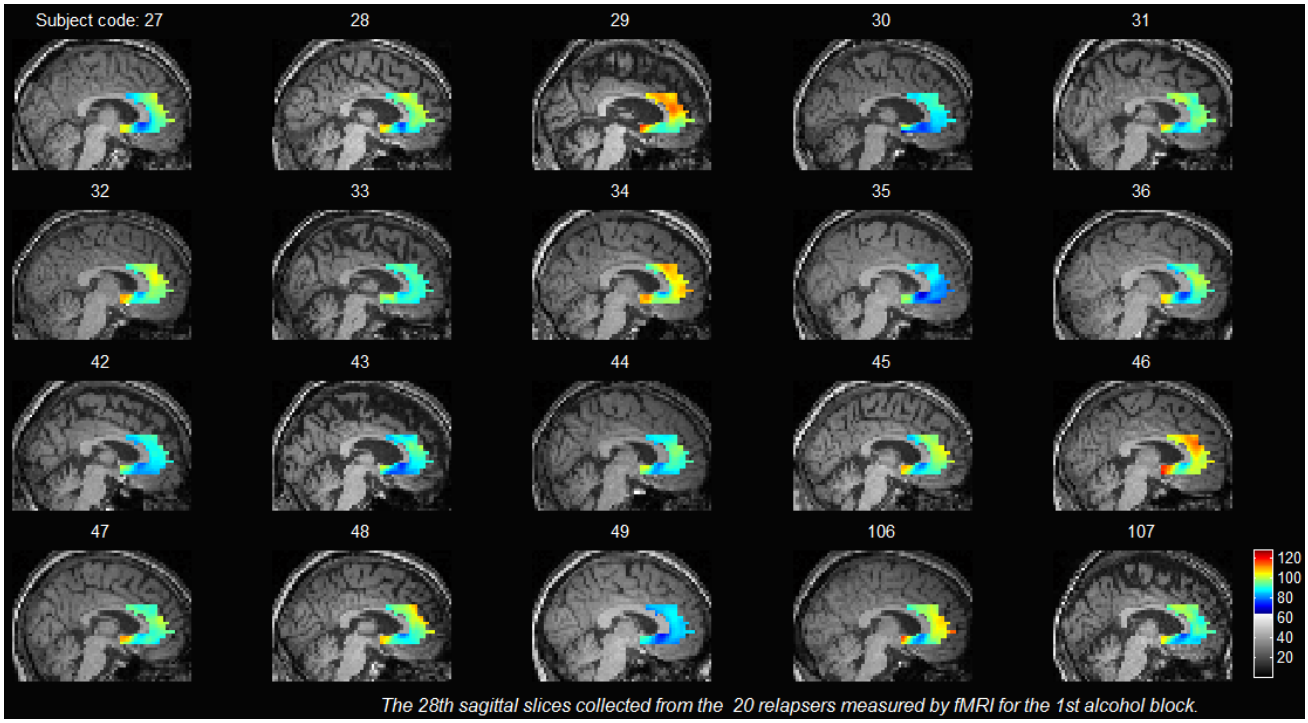


(a)

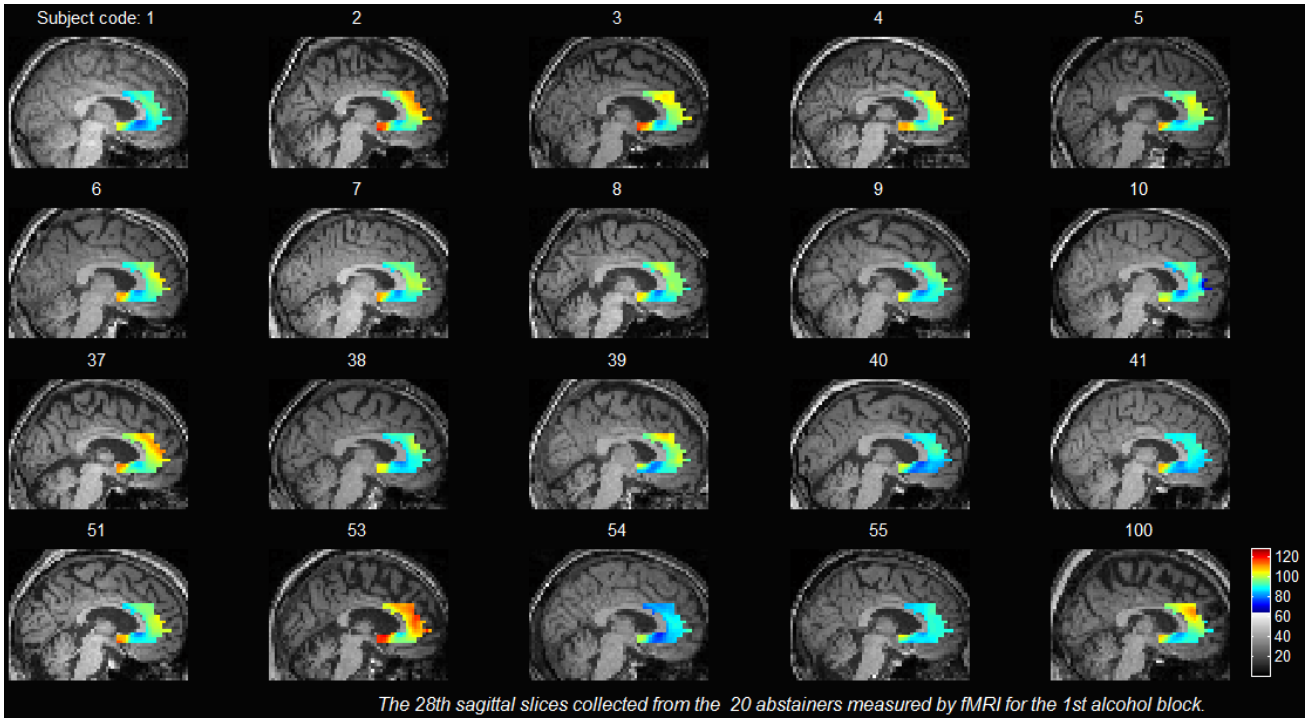


(b)

Sagittal view



(c)



(d)

AFFIDAVIT

"I, Quoc Phoi, Dam certify under penalty of perjury by my own signature that I have submitted the thesis on the topic 'Diagnostic classification and relapse prediction in alcohol dependence using fMRI, from classification algorithm to imaging approach'. I wrote this thesis independently and without assistance from third parties. I used no other aids than the listed sources and resources".

All points based literally or in spirit on publications or presentations of other authors are, as such, in proper citations (see "uniform requirements for manuscripts (URM)" the ICMJE www.icmje.org) indicated. The sections on methodology (in particular practical work, laboratory requirements, statistical processing) and results (in particular images, graphics and tables) correspond to the URM (s.o) and are answered by me. My interest in any publications to this dissertation corresponds to those that are specified in the following joint declaration with the responsible person and supervisor. All publications resulting from this thesis and which I am author correspond to the URM (see above) and I am solely responsible.

The importance of this affidavit and the criminal consequences of a false affidavit (section 156,161 of the Criminal Code) are known to me and I understand the rights and responsibilities stated therein.

Berlin, 10.01.2014 Signature

Curriculum vitae

My curriculum vitae will not be published in the electronic version of my work for privacy reasons (Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.).

ACKNOWLEDGMENTS

The creation of a dissertation is not the work of a single individual. From the bottom of my heart, I wish to express my sincere gratitude to all the people who have supported me during the implementation of this work. Without their guidance and assistance, my dissertation would not exist.

With the greatest respect, my special thanks would go to my supervisor, Prof. Dr. med. Jürgen Gallinat who has been patient to encourage and support me throughout my doctoral program. Thanks to his generous support and professional advice, I have overcome the difficulties to make this dissertation possible.

I would also like to give thanks to Dr. Mohr Johannes who has contributed a very important work in reviewing and adjusting the methodology in the first steps of this dissertation in order to make it become better. He has proved to be one of the most thorough reviewers as I expected.

A meaningful work of this dissertation was to collect the data – a work is difficult and time consuming which I have not actively participated. I would like to thank Dipl.-Psych. Katrin Charlet who has helped me get acquainted with fMRI in the first days in Berlin and also who has provided the data to perform this dissertation. Besides, I greatly appreciate the opportunity to work on such interesting data and thank Prof. Andreas Heinz, Prof. Falk Kiefer, Prof. Henrik Walter, Prof. Klaus Obermayer, Dr. Sambu Seo and Dr. Sabine Vollstädt-Klein for granting me access to the data as well as for offering the valuable suggestions for this dissertation.

I would like to thank my many friends and our kind-hearted secretary, Brigitte Butzek. They provided the constant suggestions and assistance. Without their help, I would face many difficulties while doing this dissertation. I had a great time working with them.

To my parents and my family who are always there for me, I want to give my best appreciation. Whenever thinking of them, I have gained motivation to continue to move forward, especially whenever I was discouraged.

Finally, I would like to thank Vietnamese Ministry of Education and Training (MOET) and Deutsche Akademische Austauschdienst (DAAD) for their financial support during my stay in Germany and the Bundesministerium für Bildung und Forschung (BMBF) within the Bernstein Focus: Neuronal Basis of Learning (FKZ 01GQ0911) and within the frameworks of the Nationales Genomforschungsnetz NGFN-Plus for funding the project which I worked on.