Aus dem Zentralinstitut für Seelische Gesundheit

(Direktor: Prof. Dr. med. Andreas Meyer-Lindenberg)

Klinik für Abhängiges Verhalten und Suchtmedizin

(Ärztlicher Direktor: Prof. Dr. med. Falk Kiefer)

Investigations on the influence of cue reactivity on relapse behavior in alcohol use disorder using innovative methods of statistical survival analysis

Inauguraldissertation

zur Erlangung des Doctor scientiarum humanarum

(Dr. sc. hum.)

der Medizinischen Fakultät Mannheim

der Ruprecht-Karls-Universität

zu

Heidelberg

vorgelegt von

Dipl.Math. Iris Reinhard

aus Mannheim

2021

Dekan: Prof. Dr. med. Sergij Goerdt

Referentin: Frau Prof. (apl.) Dr. sc. hum. Sabine Vollstädt-Klein

TABLE OF CONTENTS

Page

1 INTRODUCTION

In many countries, alcohol consumption is not only generally accepted, but even socially desirable. Yet the consequences of alcohol consumption for individual health, as well as economy are devastating. The recommendations for low-risk alcohol consumption of the German Federal Center for Health Education (BZgA) are less than 24g of pure alcohol for men and 12g of pure alcohol for women per day (Kalinowski & Humphreys, 2016). This corresponds to two and one standard drink/s respectively (e.g. a small glass of beer). In analogy to these guideline values, 18.1% of German adults consume alcohol in hazardous quantities (Atzendorf et al., 2019). The World Health Organization (WHO) guideline regarding brief intervention for risky drinking defines a standard drink as 10g of pure ethanol. The WHO advice is not to exceed two standard drinks per day with both men and women (Kalinowski & Humphreys, 2016).

1.1 Definition of alcohol use disorder

In the new edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5®) (American Psychiatric Association, 2013), there was a conceptual shift from the biaxial distinction between alcohol abuse and dependence to a unitary construct of alcohol use disorder (AUD) varying only in terms of severity (Dawson et al., 2013). DSM-5 defined severe AUD as an endorsement of ≥4 criteria out of 11:

- Recurrent drinking resulting in failure to fulfill role obligations
- Recurrent drinking in hazardous situations
- Continued drinking despite alcohol-related social or interpersonal problems
- Tolerance
- Withdrawal or substance use for relief/avoidance of withdrawal
- Drinking in larger amounts or for longer than intended
- Persistent desire/unsuccessful attempts to stop or reduce drinking
- Great deal of time spent obtaining, using, or recovering from alcohol
- Important activities given up/reduced because of drinking
- Continued drinking despite knowledge of physical or psychological problems caused by alcohol
- Alcohol craving

The new craving criterion is consistent with its inclusion in the International Statistical Classification of Diseases and Related Health Problems 10th Revision of the WHO (ICD-10) as one of the criteria for the diagnosis “dependence syndrome due to use of alcohol” (F10.2) (World Health Organization (WHO), 1992, 2004).

Because categorical diagnosis is still the core of the DSM-5 diagnostic system, one new approach is the RDoc initiative of the National Institute of Mental Health (NIMH) which intends to serve as a research framework for studying psychiatric disorders (Kwako et al., 2016; Lupien et al., 2017). It is based on neuroscience research and captures five domains of functioning selected on the basis of our current understanding of the neural circuitry of the brain. The domains’ fundamentals are dimensions of observable behavioral and neurobiological measures:

- Negative Valence Systems
- Positive Valence Systems
- Cognitive Systems
- Systems for Social Processes
- Arousal and Regulatory Systems

Besides the axis of the five domains there is a second axis representing the “units of analysis”. These are the different classes of variables or measures that are used to study the domains of functioning. These units of analysis include genes, molecules, cells, circuits, physiology, behavior and self-reports. Within this new framework particular constructs for craving in AUD can be investigated and used to get individualized therapy in focus. For example, the domain “Negative Valence Systems” with the subconstruct “Acute Threat (Fear)” and the motivation to relief these negative emotions with alcohol can be examined by the unit of analysis “Physiology”. An established measure is the startle reflex. On the other hand the domain “Positive Valence Systems” with the approach motivation or reward seeking can be studied by the unit of analysis “Neural Circuit” which can involve the neural cue reactivity in the ventral striatum or the orbifrontal cortex, as well as again by startle to investigate alcohol use disorder.

At the time of the studies of this thesis, DSM-IV was still in effect, but the participants met the diagnostic criteria for severe AUD in DSM-5 (American Psychiatric Association, 2010, 2013).

1.2 The cue reactivity model as a learning-based model for addictive disorders

The learning-based model of cue reactivity is based on the assumption that neutral stimuli (e.g. the sight of the favorite bar of AUD patients) which are regularly associated with alcohol intake, can trigger so-called conditioned reactions (e.g. craving as well as relapse) after long abstinence (Drummond et al., 1990). Originally neutral stimuli can be combined with the substance (unconditioned stimulus - UCS) and thus change to a conditioned stimulus (CS), which is linked to substance consumption as an unconditioned reaction (UCR). As a result, the CS alone (i.e. even without the availability of the substance) can trigger a desire for the substance in the sense of a conditioned reaction (CR) (Review see Wrase et al., 2006). Examples of conditioned stimuli (CS) are a glass of beer, the smell of a cigarette or the sight of consuming persons. In addition to these external stimuli, internal stimuli such as memories of negative experiences or experienced feelings appear as CS.

The term "cue reactivity" refers to the phenomenon that dependent individuals show physiological (e.g. neuronal activation, skin conductance), subjective (e.g., craving, joy or anxiety) or behavioral reactions (such as a relapse) in response to the presentation of substance-related stimuli. These conditioned reactions involve the phenomenon of "craving", defined as an emerging strong desire to consume the drug. Therefore a relapse often follows a confrontation with alcohol-associated stimuli.

The three-pathway psychobiological model of craving for alcohol (Verheul et al., 1999) assumes reward and relief as different craving varieties in AUD patients, whereas obsessive craving as a third type is not the focus of the studies presented here. Reward craving is induced by pleasant, positive inter- and intra-personal situations. Thus alcohol is regarded as a positive reinforcer by enhancing positive feelings and mood states associated with an opioidergic/dopaminergic dysfunction. In contrast in relief craving, triggered by negative mood or situations, alcohol acts as a negative reinforcer reducing e.g. stress and negative feelings. Relief craving is supposed to be related to GABAergic/glutamatergic dysregulation.

Based on this model Mann et al. (2009) assumed differential medication effects in AUD patients. Naltrexone as mu-opioid receptor antagonist is hypothesized to diminish the positive reinforcement mediated by alcohol in AUD individuals more suffering from reward craving. Furthermore Mann et al. (2009) hypothesized that acamprosate, interacting primarily with the glutamatergic system, reduces craving in individuals more prone to drink alcohol for relief craving.

1.3 Methods to assess cue reactivity

The principle of cue reactivity is based on the assumption that the reactions to drug-associated stimuli respectively chains of stimuli can be cognitive-symbolic (subjective craving, anxiety or pleasure), physiological (increasing heart rate, neuronal activation, skin conductance, saliva) or behavioral (drug-seeking, consuming behavior). As a result, cue reactivity can be measured using questionnaires, physiological instruments (e.g. startle, functional magnetic resonance imaging (fMRI), electroencephalogram (EEG)), or behavioral observations (e.g., assessing the days until severe relapse, drug-seeking).

Here the investigation concentrates on the physiological level and includes a neurobiological and a psychophysiological method.

1.3.1 Functional magnetic resonance imaging (fMRI)

A neurobiological method for assessing cue reactivity is the functional magnetic resonance imaging (fMRI) (Schacht et al., 2013; Yalachkov et al., 2012). FMRI is a further development of classical MRI. While classical MRI can only measure the anatomy by delineating different structures, fMRI even makes it possible to measure functions of the brain. If a brain region is activated by stimulation, there is a local increase in metabolic activity (Schad, 2002a). In exploring cue reactivity by this imaging procedure brain regions can be identified that are related to processes of conditioned reactions (Hommer, 1999). With this indirect method, changes in the local oxygenation of the blood in the brain can be detected (e.g. during sensory, motor or cognitive stimulation), which allows the identification of the functionally involved brain areas. FMRI uses this so-called Blood Oxygenation Level Dependent (BOLD) effect, which was described by Ogawa et al. (1990) and Ogawa et al. (1993) for the first time. If neuronal activity occurs in a particular brain region, for example by processing a task or by viewing an object, the metabolic activity in that region is increased. The activated area reacts with a disproportionate increase in blood flow and blood volume.

In our study fMRI data were analyzed using the neuroimaging software SPM5 (Wellcome Department of Cognitive Neurology, London, UK). Before statistical analyses can be performed, fMRI data need to be pre-processed (Schad, 2002b; Strother, 2006). The pre-processing involves removing of artefacts, correcting for head movement inside the scanner, normalization to a standard brain and smoothing to increase signal-to-noise ratio. After pre-processing SPM provides classical inference based on the general linear model by employing F- and t-tests (1st and 2nd level analysis) to identify activated brain areas. As a result, a three-dimensional statistical parametric map is obtained on which regions of cortical

activation are represented by increased parameter values. Since a large number of statistical tests are performed, namely for each voxel, a correction for multiple testing is suggested.

Previous research shows that functional magnetic resonance imaging (fMRI) is an excellent tool to explore cue-induced brain activity in AUD patients (Schacht et al., 2013). Recent studies consistently show that alcohol-related cues activate brain regions associated with motivational and emotional processing (e.g. cingulate gyrus, prefrontal areas, insular cortex, basal ganglia and amygdala) (Grüsser et al., 2004). Functional magnetic resonance imaging (fMRI) is distinguished from other functional imaging techniques, such as positron emission tomography (PET) or single photon emission computer tomography (SPECT), by non-invasiveness and lack of radiation exposure as well as a much better spatial (millimeter range) and temporal resolution (seconds range).

In our study we used the ALCUE (ALCohol Cue reactivity) paradigm (see Figure 1) to investigate the neuronal activity for alcohol-related versus neutral pictures (Vollstädt-Klein et al., 2010).

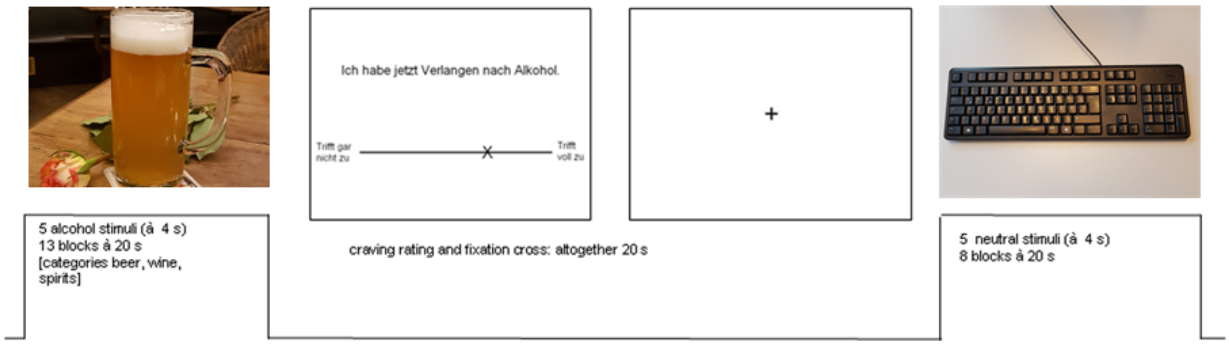


Figure 1: ALCUE paradigm

AUD patients were exposed to neutral as well as substance-related stimuli (e.g. photos of alcohol-related situations). Besides, changes in craving or substance-related behavior were measured by questionaires and structured interviews. A large number of studies over the past decade using imaging techniques have identified neural networks which are activated by substance-related stimuli in alcohol, nicotine or cocaine dependent persons (Braus et al., 2001; Brody et al., 2007; Childress et al., 1999; Franklin et al., 2007; Garavan et al., 2000; George et al., 2001; Grüsser et al., 2004; McBride et al. 2006; Schneider et al., 2001; Wrase et al., 2002). The brain regions that are typically involved in stimulus-response processes include structures that are closely associated with the mesolimbic dopaminergic network, but

also regions that are responsible for attention processes, emotion processing or inhibition: the anterior cingulate cortex (ACC), the dorsolateral prefrontal cortex (DLPFC), the orbitofrontal cortex (OFC), the insula, the nucleus accumbens (NAcc), the amygdala, the hippocampus and various occipital regions.

1.3.2 Data aggregation in fMRI

For complex statistical analysis (e.g. association of cue reactivity with relapse behavior) the aggregation of fMRI data is essential, because these extensive statistical methods are not implemented in standard whole-brain fMRI software. The statistical methods in this software are essentially based on the general linear model which is not suitable to explore relapse behavior. The different data-aggregation measures assess various aspects of neural activation, including spatial extent and intensity. Because of these variations their suitability for analyzing clinical data may differ. Standard methods calculate the mean of the contrast of parameter estimates in a predefined ROI (Poldrack, 2007). This approach is implemented in toolboxes or standalone software packages for fMRI data (i.e. MarsBar, FSL, AFNI).

1.3.3 Startle response

A psychophysiological method for the autonomous detection of the extent of cue reactivity is the measurement of the affective modulation of startle responses (Lang et al., 1990). This physiological score is obtained by the acoustic startle eye blink response. It is typically measured by electromyography (EMG) at the orbicularis oculi muscle by placing electrodes on the skin surface and is elicited by short intense bursts of acoustic white noise. By means of the amplitude of a startle response, the emotional valence of a stimulus can be estimated. Typically, unpleasant emotional stimuli increase the amplitude, while the presentation of pleasurable stimuli leads to a reduction (Bradley et al., 2006; Cook et al., 1992). For example, it is increased by fear (Vrana et al., 1988), during alcohol withdrawal (Krystal et al., 1997; Rassnick et al., 1992) or during induction of negative mood (Bradley et al., 1990), while a reduction occurs with the presentation of emotionally positive and appetitive stimuli (Lang et al., 1990). These findings qualify the startle paradigm useful to assess the valence of nonverbal, affective response towards alcohol associated stimuli in AUD patients.

Studies examining the startle response in AUD patients to understand the psychophysiological mechanisms of motivational processing that lead to craving and relapse show partially variable results. Some studies have found abstinent patients suffering from AUD to exhibit an appetitive startle response (attenuation of the amplitude) when presented with alcohol-related cues, compared to emotionally negative or negative and neutral visual stimuli (Grüsser et al., 2002; Heinz et al., 2003; Loeber et al., 2007; Mucha et al., 2000). In contrast, results of Saladin et al. (2002) point to more aversive startle amplitudes in responses to alcohol cues compared to water cues during early abstinence. Figure 2 illustrates the experimental procedure used in our study where we presented alcohol-related versus neutral, pleasant (positive) and aversive (negative) pictures. The acoustic startle eyeblink response was measured by electromyography (EMG) of the orbicularis oculi muscle. EMG data were resampled to 1000 Hz and filtered (30-500 Hz), segmented from -200 ms to 500 ms around the the startle onset, rectified, and baseline corrected (120 to -20 ms). After visual inspection, the artifact-free trials of one condition were averaged. The startle response was defined as the peak amplitude (time window 20-180 ms) of the smoothed signal (10-point moving average).

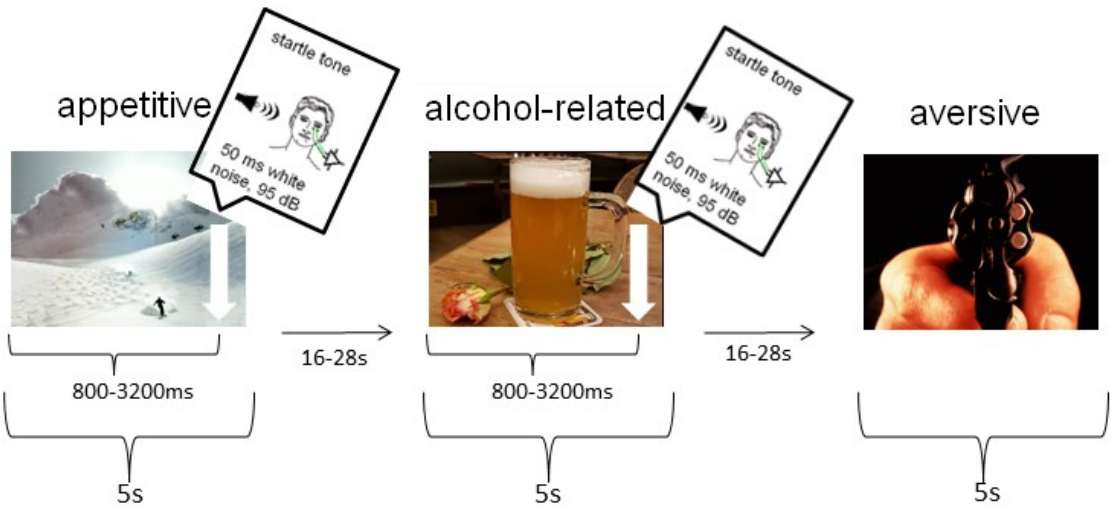


Figure 2: Startle experimental procedure

1.4 Statistical methods for relapse prediction

Survival analysis is the gold standard for statistical data analyses using time until an event occurs as outcome variable (Kalbfleisch & Prentice, 2002). Most of these analyses have to consider the problem of right-censoring which occurs when individual survival time is not exactly available because the event has not yet occurred. For identifying quantitative (and categorical as well) predictor variables for the outcome time to relapse, the Cox proportional hazards (PH) regression analysis is an analytical method for right-censored failure-time data (Cox & Oakes, 1998). The uncensored survival times are sometimes termed as event times. The Cox regression accounts for both censored and uncensored data. Two functions are often used to describe the distribution of survival times, the survivor function and the hazard function. The Cox proportional hazards model is a semiparametric model, because it assumes a parametric form for the effects of the explanatory variables, but makes no assumptions about the form of h(t) (non-parametric part of the model). The proportional hazards regression model is given by $h(t|X) = h(t) \exp(\beta_1 X_1 + \cdots + \beta_p X_p)$. The formula defines the hazard at time t as the product of two quantities, the baseline hazard function h(t) and the exponential expression exp to the linear sum of the $\beta_i X_i$ over the p explanatory X variables (Kleinbaum & Klein, 2012). The proportional hazards assumption implies that the effect of the predictors, the hazard ratio exp(β), is the same at all times t and thus is a constant, which does not depend on time t. The statistical analyses for relapse prediction were generated using SAS software, Version 9.4, of the SAS System for Windows.

1.5 Association between cue reactivity and substance use

Pilot studies with AUD patients suggest an association between drug-induced cue reactivity and substance use (Grüsser et al., 2004) as well as craving (Yalachkov et al., 2012). FMRI studies involving patients with AUD, but also social drinkers showed elevated reactions on alcohol stimuli compared to neutral stimuli in brain regions of the so-called reward system, particularly in the ventral striatum (Schacht et al., 2013; Vollstädt-Klein et al., 2010; Wrase et al., 2002). Moreover an effect of activation in the striatum and the prefrontal cortex following alcohol cues was found on the amount of alcohol in the case of a relapse (Grüsser et al., 2004).

As mentioned in chapter 1.3.3 there are heterogeneous findings and theories concerning pleasant and unpleasant effects of substance associated cues on patients suffering from AUD. Thus, on the one hand, a reduced and therefore appetitive startle response on alcohol-related stimuli was found compared to control stimuli (Grüsser et al., 2002; Heinz et al.,

2003; Loeber et al., 2007; Mucha et al., 2000), but on the other hand higher startle amplitudes could be observed, which indicate a rather aversive stimulus evaluation (Saladin et al., 2002). Compared to healthy controls, lower levels of startle response on alcohol-related cues were found for alcohol-dependent individuals (Rubio et al., 2013).

1.6 Scientific questions

The main research goal of this work is the investigation of the effect of different measures of cue reactivity on the relapse behavior in patients with AUD. So far, only few findings are documented from neurobiological and psychophysiological studies. One reason for this could be the previous use of less suitable procedures. For example, no survival analyses are implemented in standard fMRI software. When investigating the relapse behavior in a pre-defined observation period, abstinent patients cannot be assigned a relapse time (i.e., time to first severe relapse). Relapse data in alcohol-dependent patients is usually available as right-censored data, and correlation analyses are not adequate analysis methods for this type of data. As a standard method, therefore, the dichotomization of the relapse variable (relapsed versus abstinent patients) with subsequent logistic regression is often applied to investigate influencing factors. Information such as the time to relapse are not included in this kind of analyses.

In the present work complex statistical methods like survival analyses are used that involve more information and explain more variation. One main part of this thesis includes the development and examination of various methods of aggregation of whole brain fMRI data as suitable predictors of relapse time. Moreover, using modern statistical methods (e.g. calculation of the proportion of explained variation (PEV) and testing differences in PEV of various aggregation methods by means of bootstrapping methods) allow a comparison of these aggregation methods. Another main question was the investigation of the differential efficacy of drugs for relapse, also by using complex statistical methods that are new in this context.

1.6.1 Study 1

The first study of this thesis examines the alcohol-associated neuronal cue reactivity, particular in the mesolimbic reward system, such as e.g. the ventral striatum, and its influence on the time to first severe relapse or the risk of relapse of alcohol-dependent

patients with AUD. For these analyses, data from specific functionally or anatomically defined regions of the brain (ROI = region of interest) were aggregated.

The most common method calculates the mean contrast of the intra-individual parameter estimators in a predefined ROI. These and other aggregation methods capture different aspects of activation, such as the spatial extent or the intensity or both. So far, there are no studies (even beyond the addiction research) that have developed and evaluated different methods of aggregating ROI data for such investigations and compare them by using statistical characteristics and tests. In particular, the question of which type of measure is suited best as a prognostic factor for relapse is examined by complex methods of survival analyses. Therefore Cox regressions are conducted with the fMRI cue reactivity as main effect and the time to first heavy-drinking day as the failure time outcome variable, separately for each of the ROI data-aggregation measures.

In contrast to a linear regression, a Cox regression doesn't provide an explained variance. For the evaluation and comparison of the ROI data-aggregation measures, the proportion of explained variation (PEV), based on the proportional hazards model, is calculated (Heinze & Schemper, 2003) by modifying the SAS macro %RELIMPCR. Subsequently, bootstrapping methods are used for testing differences in the values of PEV, with each other and versus gold standard. These calculations are not implemented in statistical software by default and were performed after adapting a freely accessible toolbox.

1.6.2 Study 2

The second study of the thesis deals with the alcohol-associated startle reflex as a measure of cue reactivity. The modulation of the startle reflex depends on the emotional valence of the background stimuli. Therefore, the cue reactivity patterns in the startle reflex may be clearly differently pronounced in patients with AUD depending on the motivational aspects of drinking.

To the best of my knowledge, there are no studies on the suitability of the startle reflex measure as a predictor of relapse. Further, the demonstration of individually different startle reactivity patterns makes it possible to consider differential pharmacological treatment in alcohol-dependent patients according to the underlying neurobiological mechanisms (Mann et al., 2009).

In previous studies only a subset of patients benefited from the relapse-preventive effects of the substances naltrexone and acamprosate (Chick, 1995; Mann et al., 2004; Volpicelli et al., 1995). Therefore the individual differences in the alcohol-associated startle reflex and their

interaction with the medication are to be investigated (Mann et al., 2009) in the next step of study 2. In order to test the hypothesis of differential efficacy of the substances naltrexone and acamprosate under different startle reflex patterns, analyses are carried out on the basis of Cox regression models with interaction terms. Cox regression models are well established in biostatistics but are not used by default in addiction research. In particular, the modeling of interactions with drugs has not yet been used, as far as known, in the context of prediction of relapse by psychophysiological measures in AUD patients. In further analyses the continuous-by-categorical interaction is depicted by graphical representations of the log-hazard-ratio curves and the calculation of their simple slopes.

1.7 Hypotheses

1.7.1 Hypotheses Study 1

Hypothesis 1a: Cue reactivity in the mesolimbic reward system

Higher alcohol associated cue reactivity in the mesolimbic reward system is related to an increased relapse risk.

Based on learning-based models of addiction (Drummond et al., 1990), and on results from pilot studies on the association between cue reactivity and compulsory alcohol consumption (Grüsser et al., 2004 and Weiss et al.) as well as craving (Maas et al., 1998; Smolka et al., 2006), we expect cue reactivity to be related to the time to first severe relapse.

Hypothesis 1b: fMRI measures as predictors of relapse to heavy drinking

The fMRI-based aggregation measure, which is best suited as a relapse predictor, is a measure, which assesses both spatial extent and intensity of neural activation.

The most common aggregation method of the mean contrast of the intraindividual parameter estimators was strongly criticized amongst others by Poldrack (2007). This, together with an insufficient data situation to compare other aggregation measures, suggests that both aspects, spatial extent and intensity of fMRI activation, should be included in a suitable predictor.

1.7.2 Hypotheses Study 2

Hypothesis 2: Substance efficacy under various startle patterns of cue reactivity

There is a differential efficacy of the substances naltrexone and acamprosate under various startle patterns of cue reactivity. Naltrexone is supposed to be effective in patients who respond more appetitive in alcohol-cue-related startle, whereas acamprosate should reduce relapse risk in patients responding more aversive in alcohol-cue-related startle.

Studies on AUD patients with the aim to understand the psychophysiological mechanisms of motivational processing leading to craving and relapse, found inter-individual differences in the emotional evaluation of alcohol-related cues (appetitive or aversive), measured by the modulated eye blink startle (Grüsser et al., 2002; Heinz et al., 2003; Loeber et al., 2007; Mucha et al., 2000; Rubio et al. (2013); Saladin et al. (2002)).

Previous studies give evidence that only a subset of patients benefited from the relapse-preventive effects of the substances naltrexone and acamprosate (Chick, 1995; Mann et al., 2004; Volpicelli et al., 1995,). This suggests, in terms of the idea of an individualized treatment, to investigate the potential for differential pharmacological treatment due to different underlying biological mechanisms which can be identified by the affective modulation of the startle response.

2 EMPIRICAL STUDIES

This thesis involves two empirical studies. The first study was conducted to investigate cue reactivity in the mesolimbic reward system and to test hypotheses 1a and 1b. The second study addresses investigations on substance efficacy under various startle patterns of cue reactivity and tests hypothesis 2.

2.1 Study 1: A comparison of region-of-interest measures for extracting whole brain data using survival analysis in alcoholism as an example[1]

[1] Publication:

Reinhard I., Leménager T., Fauth-Bühler M., Hermann D., Hoffmann S., Heinz A., Kiefer F., Smolka M.N., Wellek S., Mann K., & Vollstädt-Klein S. (2015). A comparison of region-of-interest measures for extracting whole brain data using survival analysis in alcoholism as an example. *Journal of Neuroscience Methods, 242*, 58-64.

2.1.1 Abstract

Background

Aggregation of functional magnetic resonance imaging (fMRI) data in regions-of-interest (ROI) is required for complex statistical analyses not implemented in standard fMRI software. Different data-aggregation methods assess various aspects of neural activation, including spatial extent and intensity.

New method

In this study, conducted within the framework of the PREDICT study, we compared different aggregation measures for voxel-wise fMRI activations to be used as prognostic factors for relapse in 49 abstinent alcohol-dependent individuals in an outpatient setting using a cue-reactivity task. We compared the importance of the data-aggregation measures as prognostic factors for treatment outcomes by calculating the proportion of explained variation.

Results and comparison with existing method(s)

Relapse risk was associated with cue-induced brain activation during abstinence in the ventral striatum (VS) and in the orbitofrontal cortex (OFC). While various ROI measures proved appropriate for using fMRI cue-reactivity to predict relapse, on the descriptive level the most “important” prognostic factor was a measure defined as the sum of *t*-values exceeding an individually defined threshold. Data collected in the VS was superior to that from other regions.

Conclusions

In conclusion, it seems that fMRI cue-reactivity, especially in the VS, can be used as prognostic factor for relapse in abstinent alcohol-dependent patients. Our findings suggest that data-aggregation methods that take both spatial extent and intensity of cue-induced brain activation into account make better biomarkers for predicting relapse than methods that consider an activation’s spatial extent or intensity alone.

2.1.2 Introduction

Examining associations between functional magnetic resonance imaging (fMRI) activations and other variables (such as behavioral or genetic data, or clinical outcomes) involves the use of statistical models restricted to specific anatomically or functionally defined brain regions (regions of interest, or ROIs). A common approach is to aggregate fMRI data in these ROIs for further analyses. The most common ROI aggregation method is to calculate the mean of the contrast of parameter estimates in a predefined ROI (Poldrack, 2007). This standard approach is implemented in all toolboxes or standalone software packages such as MarsBar (Brett et al., 2002) (http://marsbar.sourceforge.net/), log_roi_batch (http://www.aimfeld.ch/neurotools/neurotools.html), ExtractVals . (http://www.umich.edu/~rcwelsh/ExtractVals.tar), Ortho (http://web.eecs.umich.edu/~hernan/Public/Programs/), FSL (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/), AFNI (http://afni.nimh.nih.gov/afni), Freesurfer (http://surfer.nmr.mgh.harvard.edu/) and rex (http://web.mit.edu/swg/rex/rex.pdf). In MarsBar's batch mode it is also possible to calculate "percent of activated voxels", i.e., the percentage of voxels in an ROI exceeding a predefined activation threshold. A similar measure is the number of activated voxels (implemented in AFNI and log_roi_batch), which can be transformed to percentage of activated voxels by dividing by the number of all voxels within the ROI and multiplying by 100. Other measures are median (implemented in FSL, AFNI, MarsBar and rex), mode (implemented in AFNI), minimum and maximum (implemented in FSL, ANI and Freesurfer). Measures aggregated with Freesurfer are not applied to all voxels within the ROI, but to all activated voxels, whereas in AFNI it is possible to consider all voxels or all non-zero voxels. FSL offers the calculation of the number of non-zero voxels within a ROI. Different ROI data-aggregation methods of fMRI data assess different aspects of fMRI brain activation, with some measuring its spatial extent, some its intensity, and some taking a combination of both into account. Because of these differences, different ROI data-aggregation methods may differ in terms of their suitability for analyzing clinical, behavioral or genetic data.

In this study, we compared the performance of different ROI data-aggregation methods, using survival analysis as an example for such analyses. The fMRI paradigm we used was a cue-reactivity task (Vollstädt-Klein et al., 2010), which we used in alcohol-dependent patients.

The incentive-sensitization model of addiction (Robinson & Berridge, 1993) suggests that addictive drugs alter mesolimbic brain networks mediating the attribution of incentive salience. Following drug use, these neural circuits may become sensitized to specific drug effects and to drug-associated stimuli such as drug-related environments, advertisements, or paraphernalia. In this way, drug-associated stimuli become salient and, through associative learning processes, can cause a pathological “wanting” to take the drug in question. These stimuli may also acquire the ability to evoke drug-like responses themselves (Drummond, 2001). The “cue-reactivity” construct comprises reactions to drug-associated stimuli on various levels, including subjective (e.g., drug craving), behavioral (e.g., drug seeking) and physiological (e.g., changes in heart rate and skin conductance) (Drummond, 2001). Pilot data suggest cue-reactivity to be associated with compulsive drug use (Grüsser et al., 2004; Weiss et al., 2001) as well as with self-reported craving intensity (Maas et al., 1998; Smolka et al., 2006), thereby highlighting the clinical significance of cue-reactivity. It is possible to examine neural cue-reactivity using fMRI (Schacht et al., 2011; Yalachkov et al., 2012).

In this study, we assessed fMRI cue-reactivity in abstinent alcohol-dependent individuals. We hypothesized that relapse risk would increase with increasing fMRI cue reactivity, especially in brain regions involved in the evaluation of cues, such as the ventral striatum (VS). To this end, we conducted survival analyses of ROI fMRI data acquired using different aggregation methods. Our main question was which of those ROI data-aggregation measures is most suitable as a prognostic factor for relapse, i.e., which aggregation measure explained the highest proportion of variation in relapse in our data when using a Cox regression.

2.1.3 Materials and Methods

Participants

Participants were a subsample of the PREDICT study, a large multicenter clinical trial (Mann et al., 2009). MR scanning was performed on 84 abstinent alcohol-dependent patients who were recruited while being treated at the Department of Addictive Behaviour and Addiction Medicine at the Central Institute of Mental Health in Mannheim, Germany. The study was approved by the Ethics Committee of the Medical Faculty Mannheim of Heidelberg University. All participants provided written informed consent according to the Declaration of Helsinki.

To be eligible for participation in the study, patients had to fulfill the following inclusion criteria: They had to have been diagnosed with alcohol dependence according to DSM-IV

and ICD-10 (as assessed using the Structured Clinical Interview (SCID I), German version (Wittchen et al., 1997)); to have abstained from alcohol for an interval between 4 and 28 days prior to inclusion in the study; to have consumed more than 14 drinks per week (for women) or more than 21 drinks per week (for men) during a continuous 30-day-period prior to hospitalization. Only right-handed subjects [handedness laterality quotient > 50, according to the Edinburgh Handedness Inventory (Oldfield, 1971)] with a normal or corrected-to-normal vision (binocular visual acuity ≥ 0.8) were included. Exclusion criteria included history of other substance abuse (apart from nicotine and cannabis), current diagnosis of psychiatric disorders requiring medication, and unstable medical conditions. Patients were included in the study upon terminating inpatient treatment. In addition to receiving standard counseling, patients were randomized either to acamprosate, naltrexone or placebo. Patients were assessed using a standardized battery of questionnaires, structured interviews, current patterns of alcohol, tobacco and drug consumption, liver parameters and sociodemographic variables (Mann et al., 2009; Mann et al., 2013). In order to allow subjects to be examined in an outpatient setting, the fMRI examination took place two weeks after the end of a three-week inpatient alcohol-detoxification program.

Treatment outcome

During a planned follow-up period of 11 weeks (80 days), patients' drinking patterns were assessed every two weeks using self-reports acquired during medical-management sessions. Self-reported drinking data were checked for plausibility by comparing them to alcohol biomarkers on the group level. No bias was found (for details, see (Mann et al., 2013)). The outcome criterion for measuring relapse during this follow-up period was defined as "days until first severe relapse", which was defined in turn as the consumption of 5 or more drinks in one day for men and 4 or more drinks in one day for women.

FMRI Paradigm

During the fMRI session, patients were presented with 15 alcohol-associated stimuli, 15 affectively neutral stimuli, and 15 neutral abstract stimuli in a block design featuring with 15 stimulation blocks (three pictures per category, 19.8 seconds per block), with stimulus blocks separated from each other by resting blocks displaying a fixation cross. Alcohol stimuli were drawn from our own picture battery, which has previously been shown to evoke craving and cue-induced mesocorticolimbic brain activation in alcohol-dependent patients (Grüsser et al.,

2000; Wrase et al., 2002). The affectively neutral stimuli were taken from the International Affective Picture System (IAPS; (Lang et al., 2008)). Pictures of all categories were comparable in terms of color distribution and contrast. Furthermore, alcohol-associated and neutral stimuli were matched for complexity. Total scan time per patient was 12 minutes and 218 volumes were acquired in total.

FMRI Methodology

Scanning was performed using a 1.5-T clinical whole-body tomography scanner (Magnetom VISION; Siemens, Erlangen, Germany) equipped with a standard quadrature head coil. We acquired 24 slices every 3.3 s (4 mm thickness, 1 mm gap) using a standard echoplanar imaging (EPI) sequence (TR=1.8 ms, TE=66 ms, α=90°) with an in-plane resolution of 64×64 pixels and a field of view (FOV) of 220×220 mm^2, resulting in a voxel size of 3.44 × 3.44 × 5 mm^3.

FMRI data were analyzed using SPM5 (Wellcome Department of Cognitive Neurology, London, UK). The first five volumes of each functional time series were discarded to eliminate T1 effects. All volumes were realigned according to the remaining first volume to correct for between-scan movements. Subjects whose head movement exceeded 3 mm were excluded. The functional images were spatially normalized to a standard EPI template in the Montreal Neurological Institute (MNI) space using a 12-parameter affine transformation with additional nonlinear components resampled with 3 × 3 × 3 mm^3 voxels. Functional data were smoothed using an isotropic Gaussian kernel (12 mm full width at half maximum [FWHM]). First-level statistical analyses were performed by modeling alcohol versus control conditions (alcohol-associated stimuli versus affectively neutral and abstract stimuli; boxcar functions convolved with the hemodynamic response function) as explanatory variables in a general linear model on a voxel-by-voxel basis. Contrast images were calculated for each subject to represent the difference in activation caused by alcohol and neutral/abstract stimuli.

Whole-brain analyses

The brain regions activated in response to alcohol cues and the effects of medication on activation will be reported elsewhere (Mann et al., submitted). In order to identify the brain regions where cue-induced activation might affect relapse, we conducted a linear correlation

analysis in SPM5 with the variables “cue-induced brain activation” (images containing contrasts between “alcohol vs. neutral/abstract”) and the dependent variable “days until first severe relapse”. Relapse-associated brain regions were identified as sets of voxels with significant linear correlations (p<.005 uncorrected for multiple comparisons, cluster size ≥ 10 voxels). In abstainers, “days until first severe relapse” are right-censored and equal the follow-up period (80 days). For this reason, the linear model we used is not statistically appropriate for our data because it treats abstinent patients as if they had relapsed. Despite this, we used it anyway to identify relevant brain regions on an exploratory basis. Thus we identified relevant ROIs using this method, after which we aggregated fMRI data in these regions with different ROI data-aggregation methods and used Cox regression to analyze this aggregated data.

ROI data-aggregation methods

For subsequent analyses examining the relationship between cue-induced brain activation and relapse, we aggregated bilateral fMRI activation data (contrast “alcohol vs. neutral/abstract cues”) from the ventral striatum (VS), the orbitofrontal cortex (OFC), and the ventral part of the anterior cingulate cortex (vACC; for a list of these ROIs, see the “Results” section). The masks we used for the VS and the OFC were probabilistic anatomical masks downloaded from the Nielsen&Hansen database (Nielsen & Hansen, 2002). For the vACC, we created our own binary mask (see Supplementary Figure 1). The masks comprise different numbers of voxels for the individual participants depending on the position of the field of view for a single measurement and also depending on potential local signal void due to susceptibility artifacts. The VS mask comprised 390-603 voxels, the vACC mask 511-556 voxels and the OFC mask 4575-5706 voxels. To compare ROIs of similar size, we additionally created a functional mask OFC_small for the OFC including only significant voxels from the group activation map (alcohol vs. neutral/abstract stimuli, p < .001, minimum cluster size 10 voxels) comprising 421-422 voxels (see Supplementary Figure 2). To aggregate activation data for the ROIs with the following different methods, we used a self-written SPM toolbox (by S.V.).

Let $P\,(v\,|r)$ be the label probability for voxel v located in ROI r. Because the probabilistic Nielsen&Hansen masks can contain non-zero values even outside of a particular anatomical ROI r, we converted these masks into binary masks with new values w_v in voxel v: $w_v = \begin{cases} 1, P\,(v\,|\,r)\ \geq\ 0.25 \\ 0, P\,(v\,|\,r)\ <\ 0.25 \end{cases}$

Let V be the set of all voxels for an individual participant (i.e., the set of measured voxels) contained in a particular binarized Nielsen&Hansen ROI in region r. The number of voxels this ROI contains is denoted by its cardinality $|V|$. Furthermore let t_v be the *t*-value in voxel v and c_v the contrast of parameter estimates in voxel v. We calculated the following five different ROI data-aggregation measures:

1. $mean_{wei}$: mean BOLD signal change, weighted with $P\ (v\ |r)$, i.e.,

$$mean_{wei} = \frac{1}{\sum_{v\,\epsilon\,V} P\ (v\ |r)} \sum_{v\,\epsilon\,V} c_v \cdot P\ (v\ |r) \qquad (1)$$

2. $mean_{unwei}$: unweighted mean BOLD signal change, i.e.,

$$mean_{unwei} = \frac{1}{|V|} \sum_{v\,\epsilon\,V} c_v \qquad (2)$$

3. Sum of *t*-values exceeding an individual threshold (see, for example, (Tost et al., 2006)) divided by number of voxels in the ROI in question. An individual threshold t_I was defined as 50% of the mean of the 5% highest *t*-values in a given individual's SPM-t-map to unbias subjects with high overall t-values. Then $norm\_sum\_t_{ind}$ is defined by

$$norm\_sum\_t_{ind} = \frac{1}{|V|} \sum_{v\,\epsilon\,V, t_v \geq t_I} t_v \qquad (3)$$

4. act_vox_{perc}: percentage of activated voxels in ROI, where voxels are considered activated if they exceed a threshold of p<.05, uncorrected. A liberal threshold of p<.05 was chosen to yield larger variance. With stricter thresholds more zero percentages would occur. Let $t_{.05}$ be the critical *t*-value corresponding to the significance level of α <.05 (one-sided). Then

$$act\_vox_{perc} = \frac{1}{|V|} |\{v\,\epsilon\,V \mid t_v \geq t_{.05}\}| \cdot 100 \qquad (4)$$

5. max: maximum signal change in ROI, i.e.,

$$max = maximum(\{c_v \mid v\,\epsilon\,V\}) \qquad (5)$$

Survival analysis and comparison of importance of the ROI data-aggregation methods

Using Cox regressions conducted in the software package SAS (Version 9.2, SAS Institute Inc., Cary, NC, USA), we examined the association between fMRI cue-reactivity – in the VS, OFC as well as the vACC – and time to first severe relapse in the follow-up-period. We conducted this analysis separately for each of the five ROI data-aggregation measures listed in section 1.6. Furthermore, we evaluated and compared the relative importance of these different ROI data-aggregation measures as prognostic factors for treatment outcomes. To do this, we calculated the proportion of explained variation (PEV) (Heinze & Schemper, 2003) by means of adapting the SAS macro %RELIMPCR. PEV is based on the measure V, originally proposed by Schemper and Henderson (Schemper & Henderson, 2000). Differences in the values of PEV for the various ROI data-aggregation methods were statistically tested using a bootstrap resampling scheme (Heinze & Schemper, 2003).

2.1.4 Results

Patient characteristics

We excluded data from 9 participants from our statistical analysis due to heavy head movement (> 3mm in either the *x*, *y* or *z* direction or > 3° rotation in any direction) during the scanner session. We excluded 5 participants due to technical problems, brain abnormalities or incompliance. Another 21 patients could not be examined because they relapsed before their fMRI appointment. The final sample consisted of 49 patients (42 smokers, 39 males, aged 44 ± 9 years, range 27–63 years). Before examination, patients had been abstinent for 37 ± 5 days (range 29–64 days), had been alcohol dependent for 14 ± 9 years (range 0–34 years), and had consumed 151 ± 83 g alcohol per day immediately prior to detoxification (range 10–334g per day). Patients lasted 54 ± 27 days (range 2–80) until their first severe relapse. During this period, *N*=20 of the *N*=49 patients suffered a severe relapse.

Explorative fMRI analyses

Using the linear regression specified in section 1.5, we identified the VS [(*x,y,z*) = (12,3,-6), *t*=3.42, 32 voxels], the OFC [(*x,y,z*) = (-24,42,-12), *t*=2.94, 11 voxels], and the vACC [(*x,y,z*) = (-3,36,-9), *t*=3.23, 14 voxels] as regions whose activation was negatively linearly associated with time to first severe relapse (N.B.: these data are right-censored for abstainers; for details, see section 2.1.3). In other words, we found increased post-treatment cue-induced

brain activation to be associated with increased relapse risk in the VS, the OFC, and the vACC. Figure 1 displays this association.

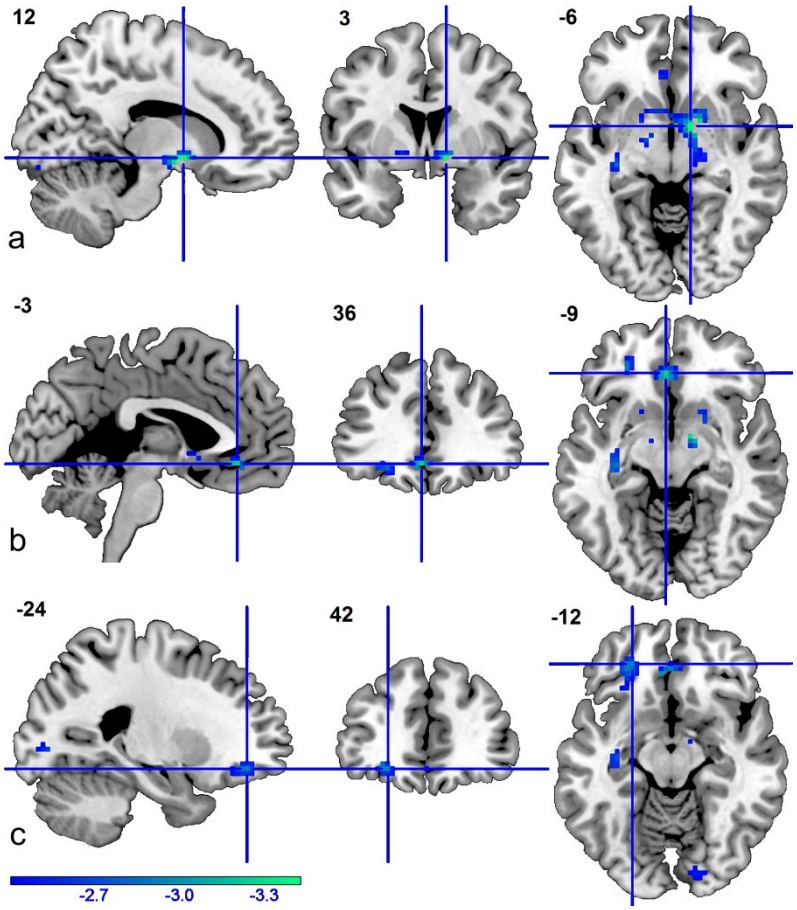


Figure 1: Negative association between fMRI cue reactivity and time to first severe relapse in (a) the ventral striatum (VS), (b) the ventral anterior cingulate cortex (vACC), (c) the orbitofrontal cortex (OFC), p<.01 uncorrected for illustration purposes

Survival analyses

In our Cox regressions, we found several ROI data-aggregation measures to have significant effects on time to first severe relapse, with hazard ratios indicating higher relapse risk with higher cue-reactivity (see Table 1). For the VS, all measures except $\max$ were found to be

significant prognostic factors, whereas for the OFC, only the measures $\mathrm{norm\_sum\_t_{ind}}$ and $\mathrm{act\_vox_{perc}}$ showed any trend (p<.1). For the OFC_small a trend (p=.0846) for measure max could be demonstrated. For the vACC, none of the measures showed significant effects. For illustration purposes, Figure 2 displays the Kaplan-Meier estimates of the survival rates, comparing patients with high cue-reactivity in the VS to those with low VS cue-reactivity (after median split and using $\mathrm{norm\_sum\_t_{ind}}$ as an independent variable).

Table 1: Association between BOLD signal change and days until 1st severe relapse: hazard ratio, p-value and proportion of explained variation (PEV marginal) for a single ROI measure in a Cox regression

	Association between cue-reactivity and days until 1st severe relapse			
	VS	vACC	OFC	OFC_small
	Hazard ratio / p-value / PEV	Hazard ratio / p-value / PEV	Hazard ratio / p-value / PEV	Hazard ratio / p-value / PEV
$\mathrm{mean_{wei}}$	2.77 / 0.0256 / 6.80%	not applicable because of binary mask	2.51 / 0.1377 / 3.24%	not applicable because of binary mask
$\mathrm{mean_{unwei}}$	2.63 / 0.0327 / 6,24%	0.98 / 0.9679 / 0%	2.30 / 0.1943 / 2.46%	1.05 / 0.9127 / 0.03%
$\mathrm{norm\_sum\_t_{ind}}$	2.78 / 0.0011 / 11.44%	0.64 / 0.3802 / 0.91%	2.13 / 0.0973 / 4.03%	0.98 / 0.9436 / 0.02%
$\mathrm{act\_vox_{perc}}$	1.03 / 0.0024 / 9.98%	0.99 / 0.4014 / 0.93%	1.02 / 0.0639 / 4.54%	1.00 / 0.8671 / 0.04%
max	1.26 / 0.2446 / 2.19%	0.82 / 0.4083 / 0.57%	0.76 / 0.1989 / 2.20%	0.63 / 0.0846 / 4.10%

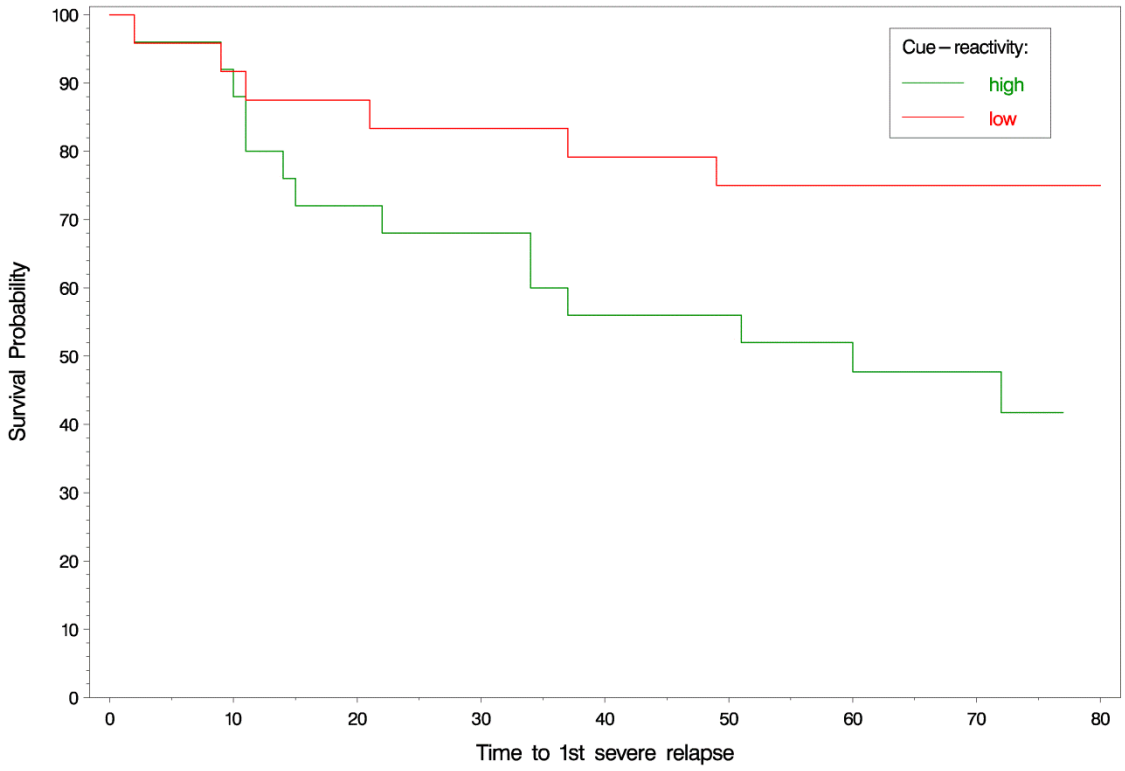


Figure 2:

Association between BOLD signal change in the VS and days until 1st severe relapse: Kaplan-Meier estimates of survival rates in patients with low vs. high cue-reactivity (measure $\mathrm{norm\_sum\_t_{ind}}$ after median-split for illustration purposes)

Evidence suggests that the ROI data-aggregation measures under examination differ in terms of the amount of explained variation in predicting relapse by fMRI cue-reactivity (see Table 1). Measure $\mathrm{norm\_sum\_t_{ind}}$ explained the highest proportion of variation in the VS, coming in just ahead of $\mathrm{act\_vox_{perc}}$. In the OFC, however, the measure $\mathrm{act\_vox_{perc}}$ explained more variation than $\mathrm{norm\_sum\_t_{ind}}$. Testing the differences in marginal PEV statistically (without adjusting for multiple comparisons) we found $norm\_sum\_t_{ind}$ to have a significantly

higher PEV than $\max$ in the VS (p = 0.0305). The remaining differences in PEV between measures in the VS did not appear to be significant. For the OFC and the OFC_small, none of the ROI data-aggregation methods showed any significant advantage over the others in terms of a significantly higher PEV. Comparing the PEV for each ROI data-aggregation measure between the different regions indicated that the measures $norm\_sum\_t_{ind}$ (p=.0126), $\mathrm{act\_vox_{perc}}$ (p=.0113) and $\mathrm{mean_{unwei}}$ (trend: p=.0577) were superior in the VS compared to the OFC_small. Furthermore, a trend for the difference between $norm\_sum\_t_{ind}$ in VS versus OFC could be revealed (p = .0878).

2.1.5 Discussion

In our study, we were able to show fMRI cue-reactivity in the VS to be suitable as prognostic factor for relapse in abstinent alcohol-dependent individuals in an outpatient setting. Data aggregated from the OFC and the OFC_small were suitable at a statistical trend level. In line with our hypotheses, participants with increased neural cue-reactivity demonstrated increased relapse risk. As expected, the ROI data-aggregation measures we used differed in terms of the proportion of variation in relapse behavior they explained, and thus in terms of their usefulness in predicting relapse with fMRI cue-reactivity.

Two of the brain regions (VS and OFC) that our exploratory pre-analysis identified are known to be involved in reward valuation (Schacht et al., 2013). Furthermore, brain structure and function in these regions have been shown to be associated with relapse (Beck et al., 2012). VS and OFC cue-reactivity would thus seem to be appropriate biomarkers for use in testing the suitability of different ROI aggregation measures for predicting relapse in alcoholism.

We found that the different ROI data-aggregation measures showed the largest differences in predictive capacity in the VS. Additionally, for the VS we found the measure $\mathrm{norm_{sum_{t_{ind}}}}$, which combines *t*-values exceeding an individually defined threshold, to be a more important prognostic factor for predicting relapse compared to measure max, as it explained more variation in relapse behavior. Descriptively, this measure was also superior compared to all other measures. This finding indicates that a method combining the spatial extent of cue-induced brain activation with the intensity of this activation is a more appropriate biomarker for predicting relapse than measures that only take into account an activation's spatial extent (such as the percentage of activated voxels $\mathrm{act\_vox_{perc}}$) or intensity (such as the mean value $\mathrm{mean_{wei}}$). Additionally, as expected, taking only one voxel intensity value into account per patient ($\max$) yielded the poorest predictive capacity.

To our knowledge, there exist only three methodological papers or published abstracts on the topic of aggregation measures for fMRI data from ROIs. Poldrack (Poldrack, 2007) describes ROI aggregation as a general method for measuring brain activation in ROIs that aims to facilitate data exploration, reduce the number of statistical tests necessary, and limit analyses to functionally defined regions. Furthermore, he argues that simply considering the mean activation within a ROI is problematic because if "there exist areas of both activation and deactivation within the region, these may cancel each other out," thereby making this approach suitable only for small or functionally homogenous regions. Cohen & DuBois have suggested that voxel-counting, as performed in measure act_vox_{perc}, might be less reliable as a measure of activation than to measures of signal change such as $\mathrm{mean}_{\mathrm{wei}}$. In contrast, one recently published abstract (Bhattacharyya et al., 2013) finds act_vox_{perc} to be especially suitable and more robust than the measure max in populations with high inter-subject variation in noise. This latter finding is in line with our findings.

On a descriptive level, we found neuronal cue-reactivity data extracted from the VS to have higher prognostic value than that taken from either the OFC or the vACC in terms of the amount of variation in relapse behavior they explain. One reason for this could be the relatively small size of the VS: larger regions may also contain voxels that are not specifically involved in cue processing and relapse. Aggregating estimates from such large regions may decrease effect sizes simply due to averaging. To examine this explanation, we used the mask OFC_small, which contains only activated voxels within the OFC and is of comparable size as the VS and vACC. However, the VS data are also superior in predicting relapse compared to the OFC_small data. A reason for this finding might be that OFC_small does not contain the relevant voxels for relapse prediction. Anyway, the definition of the ROIs is very crucial for the analysis. One has to compromise between small ROIs, which might contain not enough voxels with relevant information, on the one hand, and the use of large masks, for which effects might decrease because of averaging, on the other hand.

VS, OFC and vACC cue reactivity are all associated with each other (Yalachkov et al., 2012). The OFC and the vACC (which is also known as the subgenual part of the anterior cingulate cortex) are involved in producing affective states. The VS and the OFC play prominent roles in valuing reward outcomes. Furthermore, the VS is involved in reward anticipation (Diekhof et al., 2012). Additionally, the VS plays a key role in mesolimbic pathways, where it provides a functional interface between the limbic and motor systems, integrating signals with emotional content and motivational significance into outputs to brain regions responsible for motivational behavior (Groenewegen & Trimble, 2007). The significant role the VS plays in

these different processes that are all important for the maintenance of alcohol dependence may help explain its superiority to other brain regions in predicting relapse.

This is the first study examining the suitability of different ROI data-aggregation measures as prognostic factors for relapse in alcohol dependence. Because ROI aggregations are used in many other fields of research, our results are applicable outside of addiction research. There are many fields of medicine that are concerned with relapse prediction, with examples including autoimmune-disease relapse, recurring psychiatric episodes, and cancer relapse. Relapse predictors might be fMRI patterns in relapsing–remitting multiple sclerosis patients, which were shown to be associated with subtle performance alterations (Wojtowicz et al., 2014). Emotional reactivity in patients with recurrent depression was already shown to be able to predict relapse (Farb et al., 2011). Another example are fMRI alterations in the early disease relapse of gynecological cancer, which might not be detectable as morphological changes (Alvarez Moreno et al., 2012). Furthermore, it is necessary to aggregate fMRI data into ROIs when conducting complex statistical analyses that are not implemented in standard fMRI software using fMRI data as variables of interest. Another application could be in the analysis of data from genome-wide association studies (GWAS) and their association with fMRI data. GWAS data sets are large and therefore cannot be analyzed voxel-wise due both to limited computational resources as well as multiple-comparison issues.

Limitations

In classical linear regression models for normally distributed response variables, R^2 is the standard measure of predictive accuracy. For Cox regression however, the biostatistical literature contains several different measures intended to accomplish the same purpose (for a recent overview, see (Bøvelstad & Borgan, 2011)). Unfortunately, there is no consensus in the literature as to which of these options should be implemented in practice. The proportion of variation of survival times that can be explained using prognostic factors is very low. It typically ranges from 10–35% in survival studies (Heinze & Schemper, 2003).

The dropout rate in our study was very high (42%) and included both fMRI dropouts and relapse cases. The conclusions of our study may therefore only be valid for a subgroup of patients, namely those who do not relapse directly after inpatient treatment.

Although, we found descriptive differences in PEV between the different measures and the different ROIs, only one difference was statistically significant: Measure $\mathrm{norm\_sum\_t_{ind}}$ was superior to the measure max for the VS. Differences between ROIs only reached significance

for the VS compared to OFC_small. A statistical trend was observed for VS compared to OFC. Hence, the results should be interpreted with caution and need to be replicated in future studies involving a greater sample size.

Conclusions

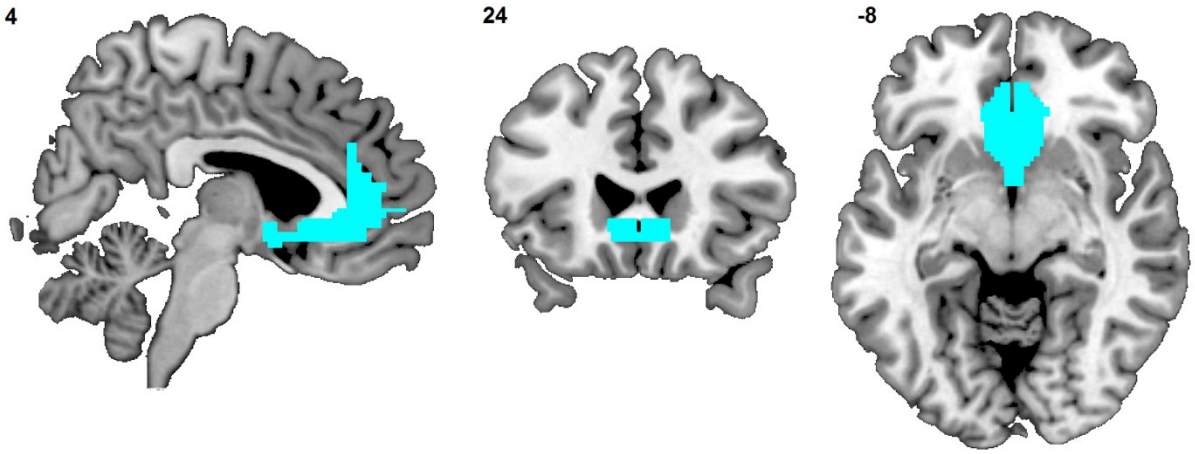
Our results demonstrate that fMRI cue-reactivity, especially in the VS, can be used as prognostic factor for relapse in abstinent alcohol-dependent patients. The measure $norm\_sum\_t_{ind}$ proved to be the most important prognostic factor for the VS, at least at a descriptive level, which suggests that a measure combining both the spatial extent and the intensity of cue-induced brain activation might predict relapse better than measures considering either spatial extent or activation intensity alone. As a consequence, it seems that fMRI cue-reactivity, especially when measured in the VS, can be used as prognostic factor for relapse in alcohol dependence.

Conflict of interest

Falk Kiefer received honoraria from Desitin and Lundbeck. Karl Mann received research grants from Alkermes, Lundbeck and Merck Sharp & Dohme. He also works as a consultant for Lundbeck. The other authors have no conflicts of interests to declare.

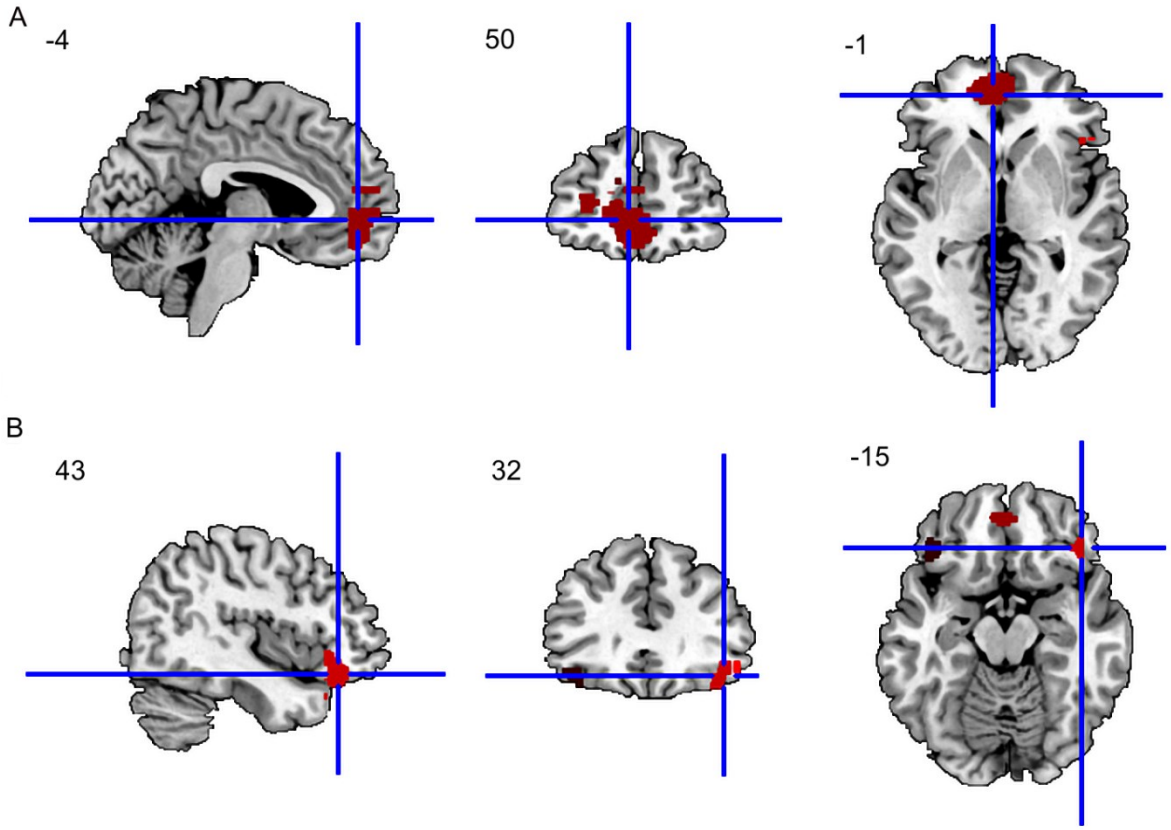
Acknowledgements

This study was supported by the German Federal Ministry of Education and Research (BMBF, Grant #01 EB 0410). Merck-Lipha and Bristol-Myers Squibb both donated medication for use in the study. We would like to thank Bernhard Croissant, Alexander Diehl, Alexander Georgi, Oliver Klein, Rosemarie Krämer, Jörg Nikitopolous, Anke Zimmer-Butz and the PREDICT Study research team for their assistance in data collection.



Supplementary Figure 1:

Binary mask for the ventral anterior cingulate cortex (vACC)



Supplementary Figure 2:

Functional mask OFC_small for the orbitofrontal cortex, A: medial part, B: lateral part

2.2 Study 2: Association between alcohol-cue modulated startle reactions and drinking behavior in alcohol dependent patients - results of the PREDICT study[2]

[2] Publication:

Leménager T.*, Hill H.*, Reinhard I.*, Hoffmann S., Zimmermann U.S., Hermann D., Smolka M.N., Kiefer F., Vollstädt-Klein S., Heinz A., & Mann K. (2014). Association between alcohol-cue modulated startle reactions and drinking behaviour in alcohol dependent patients - results of the PREDICT study. *International Journal of Psychophysiology, 94,* 263-271.

*T. Leménager T., H. Hill and I. Reinhard contributed equally to this manuscript.

2.2.1 Abstract

Previous research on alcohol dependent patients has shown that variations in eyeblink startle response can be used as an indicator of their emotional responses to alcohol-related stimuli. Postulating that reactions on substance associated stimuli are controlled by either a negative or a positive affective processing system, we expect that abstinent alcoholics react differently (within-group) in the emotional evaluation of alcohol cues. Furthermore, we assumed the startle response to covary with medication response to acamprosate and naltrexone.

We measured 74 detoxified inpatients' acoustic startle responses while they were being presented with alcohol-related images as well as affectively negative, neutral, and positive pictures before they were randomized to pharmacotherapy.

Group-mean startle peak amplitudes were lowest for alcohol-related cues. The relative startle response (alcohol cues set in relation to the other stimulus categories) did not correlate with craving for alcohol (OCDS) or alcohol cue induced self-ratings of arousal, valence and craving. Patients with a lower percentage of abstinent days in the 90 days prior to the last drinking day showed a lower ("more appetitive") startle response to alcohol cues. A survival analysis using the time to first heavy drinking day as the survival criterion revealed a significant interaction between alcohol-cue startle responses and medication type.

The results indicate that the psycho-physiological measure of emotional evaluation of alcohol cues includes unconscious processing not reflected by conscious self-ratings. Furthermore, our result of a differential medication effect may encourage further studies to use biological characteristics to stratify patients as a step towards individualized treatment for alcohol dependence.

Keywords: alcohol-cue modulated startle response, alcohol dependence, individualized treatment.

2.2.2 Introduction

The development and maintenance of drug addiction are regarded as a repetitive excessive behaviour which increasingly turns into an automatized action, difficult to control intentionally. This automatic behaviour is reinforced by conditioned learning processes, associated with neurobiological, psychophysiological and psychological alterations (Robinson and Berridge, 2013; Everitt and Robbins, 2013). One method which investigates these underlying

mechanisms by assessing various reactions elicited by drug-associated cues is designated as “cue reactivity”. Cue reactivity to alcohol-related cues has been studied extensively in alcohol-dependent patients in an effort to increase our understanding of craving and relapse (Drummond et al., 1990; Hammersley, 1992; Loeber et al., 2009; Niaura et al., 1988; Rohsenow et al., 1990).

Drummond et al. (1990) employed a cue-reactivity paradigm in which cue reactivity encompassed three types of responses to drug-associated cues: (1) autonomic (e.g., modulated startle response, skin conductance), (2) cognitive-symbolic (e.g., subjective craving) and (3) behavioural (e.g., drug-seeking behaviour) (Drummond et al., 1995). The relation between these three components of addicted patients’ cue reactivity is controversial (Drummond, 2001; Gruesser et al., 2002; Heinz et al., 2003; Tiffany and Conklin, 2000). Although several studies on cue reactivity suggest that drug cues consistently lead to increased craving and changes to indicators of autonomic reactivity, such as heart rate, sweat gland activity, temperature and salivation (Niaura et al., 1988; Rohsenow et al., 1990, Drummond et al. 1990, Glautier et al., 1994), the literature has only found there to be a modest relationship between cue-induced craving and autonomic reactions (Tiffany et al., 2000; Drummond, 2001).

One psychophysiological method that can be used to assess autonomic reactions to drug-associated stimuli is observing the affective modulation of startle responses (Lang et al. 1990). The amplitude of a startle response provides a way of estimating the emotional valence of a stimulus. Presenting subjects with unpleasant emotional stimuli typically increases the amplitude of their startle response, while presenting them with pleasant stimuli decreases it (Bradley et al., 2006; Cook et al., 1992).

Studies, examining the startle response in alcohol-dependent patients in order to understand the psychophysiological mechanisms of motivational processing leading to craving and relapse, indicate partially discrepant results. Four studies have found abstinent alcohol-dependent patients to exhibit an appetitive startle response (attenuation of the amplitude) when presented with alcohol-related cues, compared to emotionally negative or negative and neutral visual stimuli (Gruesser et al., 2002; Heinz et al., 2003; Loeber et al., 2007; Mucha et al., 2000). Furthermore, Rubio et al. (2013) assessed differences in startle responses between alcohol-dependent and healthy subjects and also observed the alcohol-dependent group to exhibit “appetitive” amplitudes in response to alcohol cues. Additionally, Loeber et al. (2009; 2007) observed appetitive startle responses to be stable over 3 weeks during which patients underwent an inpatient medical detoxification program. Appetitive responses

were also observed in subjects at genetic risk for alcoholism (Zimmermann et al., 2004). In contrast to these studies, Saladin et al. (2002) found alcohol-dependent patients to have higher startle amplitudes in responses to alcohol cues than to water cues during early abstinence, indicating that alcohol cues had a more aversive connotation for them.

In accordance with the finding that cue-induced craving is only weakly correlated with other cue-induced autonomic reactions (Tiffany et al., 2000; Drummond, 2001), startle response to alcohol cues was not found to be linked to either subjective craving or patients' valence ratings of alcohol stimuli (Gruesser et al., 2002). Based on these results, we suggested that future studies consider different individual cue-reactivity patterns in alcohol-dependent patients (Heinz et al., 2003; Mann et al., 2009).

Therefore, we investigated inter-individual differences in the emotional evaluation of alcohol-related cues (appetitive or aversive), measured by the modulated eye blink startle response in abstinent alcohol dependent patients. We assessed the association between these startle reactions and self-reports of craving. Furthermore, we investigated if alcohol cue self-ratings, assessing valence, arousal and subjective craving, will explain a part of the startle-reactions.

The assumed different patterns of startle response may further hold the potential for differential pharmacological treatment due to different underlying biological mechanisms. Studies assessing the efficacy of naltrexone and acamprosate for treating alcohol dependence indicate that not all patients benefit from these treatments (Chick, 1995; Volpicelli et al., 1995). According to the hypotheses of the PREDICT study outlined in Mann et al. (2009), we investigated individual differences in abstinent alcohol-dependent patients' startle response to alcohol-related cues and their interaction with medication. These assumptions of differential medication effects were deduced from the 'three-pathway psychobiological model of craving for alcohol' (Verheul, 1999), which postulated different craving varieties of alcohol dependent patients. The hypotheses of the PREDICT study focused on the two most empirically convincing craving varieties, i.e. reward and relief craving (Heilig et al., 2011; Gloeckner-Rist et al., 2013; Mann et al., 2013). Reward craving is provoked by pleasant, positive inter- and intra-personal situations, such as positive mood states and agreeable social events. In this craving variety, alcohol acts as a positive reinforcer, inducing or enhancing positive feelings and mood states that are associated with an opioidergic/dopaminergic dysfunction, most marked in the striatum. Since naltrexone blocks mu-opioid-receptors (Hoelter and Spanagel, 1999), we assumed that it diminishes the positive reinforcement mediated by alcohol. Based on this explanation of the reward craving variety, we assumed that the more appetitive the patients responded in the alcohol-cue-

related startle, the more they link the effect of alcohol to positive reinforcement and the more naltrexone should be effective in decreasing the relapse risk.

In contrast, relief craving is postulated to be triggered by negative mood or situations, inducing withdrawal relief states. In such settings, alcohol is consumed in order to reduce either stress, negative feelings or withdrawal symptoms. Thus, alcohol acts as a negative reinforcer related to GABAergic/glutamatergic dysregulation, especially in the amygdala and hippocampus. Since acamprosate interacts primarily with the glutamatergic system, it is supposed that the more aversive the patients responded in the alcohol-cue-related startle, the more acamprosate should reduce relief craving and relapse risk (Kotlinska and Bochenski, 2008; Schwartz et al., 2010).

2.2.3 Materials and Methods

Subjects

Data were taken from 74 detoxified abstinent inpatients (Table 1) who fulfilled diagnostic criteria for alcohol dependence according to the DSM-IV (American Psychiatric Association, 2000) and the ICD-10 (World Health Organization (WHO), 1992). All participants signed an informed consent statement that had been approved by the Ethics Committee of the Mannheim Medical Faculty of the University of Heidelberg. According to the exclusion criteria of the PREDICT study, patients had no other axis-I psychiatric disorders and no other substance misuse and dependencies (apart from tobacco and cannabis), nor were they using any substances harmfully at the time point of study onset. Three patients had a current cannabis misuse, but were abstained a minimum of two weeks due to their inpatient treatment. Regarding personality disorders, one patient was diagnosed with an anancastic personality disorder, one had a dissocial- and three patients suffered from a borderline personality disorder.

We made allowances for unproblematic use of caffeine and cannabis (urine drug testing). In order to assess psychiatric comorbidities, the SCID I and II were performed by trained researchers. Furthermore, all subjects were free of all psychotropic medication and also had not received any benzodiazepines or clomethiazol for at least 7 days before randomization. Patients with severe neurological or hepatic impairment (e.g. liver cirrhosis) were excluded.

Baseline psychometric and psychophysiological data were collected at the end of a two-to-three-week inpatient treatment period concluding just prior to the onset of the outpatient

study. A second measurement was conducted three weeks later. By this time, patients had already received medication (acamprosate (n=27), naltrexone (n=35) and placebo (n=12)) for about two weeks. They were randomized to acamprosate, naltrexone or placebo in proportions 2: 2: 1 using an imbalanced block-randomization algorithm within the framework of the PREDICT study. Patients were administered pharmacotherapy for 12 weeks. The three medication groups did not differ significantly in age, gender, age of onset, drinking patterns, severity of dependence (ADS), depressive symptoms (BDI) and the duration of abstinence before the startle assessment (Table 1). The average of the BDI scores in the sample was 6.6 and can be considered as normal (Hautzinger et al., 1995). All patients received “Medical Management” counselling. A detailed description of this study’s design and overall results can be found in Mann et al. (2009¸ 2013).

Table 1

Description of the sample (mean ± SD).

	BMBF (no Startle)	STARTLE	p-value	Naltrexone	Acamprosate	Placebo	p-value
Sample size	352	74		35	27	12	
Gender (male)	277 (78.7%)	50 (67.6%)	.039	22 (62.9)	20 (74.1)	8 (66.7)	.644[Chi]
Age	45.6 ± 8.7	43.7 ± 8.7	.093[t]	43.7 ± 8.2	42.3 ± 8.4	46.9 ± 10.4	.319[F]
Age at onset of alcoholism	30.9 ± 10.4	31.0 ± 9.4	.937[t]	31.5 ± 9.4	31.0 ± 9.7	29.2 ± 9.2	.757[F]
Alcohol Dependence Scale (ADS)	15.0 ± 6.7	15.1 ± 7.0	.897[t]	14.1 ± 6.0	16.6 ± 7.1	14.9 ± 9.5	.378[F]
Obsessive Compulsive Drinking Scale	13.6 ± 6.1	13.8 ± 6.2	.754[t]	13.9 ± 6.3	13.3 ± 6.7	14.9 ± 4.7	.747[F]
Percent heavy drinking days (Form 90)	63.7 ± 23.5	60.7 ± 23.1	.223[M]	59.6 ± 22.3	59.5 ± 25.2	66.5 ± 20.9	.635[K]

	BMBF (no Startle)	STARTLE	p-value	Naltrexone	Acamprosate	Placebo	p-value
Sample size	352	74		35	27	12	
Percent drinking days (Form 90)	65.8 ± 22.3	62.8 ± 22.5	.201[M]	61.7 ± 22.0	61.8 ± 24.1	68.4 ± 21.4	.507[K]
Number of standard drinks per drinking day (Form 90)	18.7 ± 10.8	20.1 ± 16.8	.483[t]	23.1 ± 17.7	18.9 ± 18.1	18.0 ± 7.0	.518[F]
Duration of abstinence prior to the last drinking day [days]	22.3 ± 4.4	21.2 ± 4.3	.054[t]	21.6 ± 4.1	21.4 ± 4.3	19.9 ± 4.8	.503[F]
Beck Depression Inventory (BDI)	6.6 ± 5.9	6.6 ± 6.1	.961[t]	6.9 ± 6.4	6.1 ± 6.4	7.0 ± 5.0	.873[F]

F: Anova; Chi: Chi-Square; K: Kruskal-Wallis-Test; t: t-Test; M: Mann-Whitney-U-Test; One standard drink=12 grams; SD: Standard Deviation; Form 90: the time frame used to compute this figure was the 90 days prior to the last drinking day.

Psychometric testing

We assessed craving using the Obsessive Compulsive Drinking Scale (OCDS, Anton et al., 1995; Mann and Ackermann, 2000). Furthermore, patients' drinking patterns during the 90 days prior to their last drinking day were assessed using the Form 90 interview (Miller, 1996). Patients completed the OCDS questionnaire and Form 90 before their startle response was measured.

Assessment of relapse

Main endpoints were the time to the first heavy-drinking day (defined as 5 drinks (60 g alcohol) or more drinks per day for men and 4 (48 g) or more drinks per day for women), as well as patients' abstinence status 90 days after randomization. Patients who did not return for follow-up sessions were assumed to have resumed heavy drinking on the day of their missed session.

Psychophysiological assessment

Stimulation: Twelve images depicting alcoholic beverages were used as well as 12 neutral, 12 aversive (negative valence) and 12 pleasant (positive valence) pictures taken from the International Affective Picture System (Lang et al., 1999). Pictures were presented for 5 s on a 15-inch monitor that was placed at a distance of 0.5 m in front of the subject using ERTS software (BeriSoft, Frankfurt, Germany). The startle probe was a 50 ms burst of white noise (95 db, instantaneous rise and fall), delivered binaurally by means of On-Ear headphones and an external HiFi amplifier (Kenwood KAF-110) connected with the computer soundcard (Soundblaster Pro) to the computer at a randomly chosen interval between 800 and 3200 ms (average 2000 ms) after picture onset. Loudness of the startle probe was calibrated using an artificial head and a sound level meter before the study. The startle probe was applied for eight of the 12 pictures in each category. In addition, eight startles without a picture cue were presented randomly. The inter-trial intervals varied from 16 to 28 s (average 23.4 s). For approximately the first half of the subjects, we used individually randomized stimuli sequences. For the remaining subjects, we alternated between two versions of previously randomly arranged stimuli. Before the experiment, each subject performed a practice block with five picture-startle trials in order to prevent the habituation that occurs during the first few

startle stimuli biasing the measured effect of the cues (Lang et al., 1998). After the main experiment, the 48 picture-stimuli were presented again and the patients were asked to rate the pictures on a visual analogue scale (VAS) for emotional valence (ranging from positive to negative), affective arousal (tranquil to exited) and subjective stimuli induced alcohol craving (high to low). The VAS was a paper-pencil version with a length of 100 mm.

Patients' startle response was recorded as the electromyographic (EMG) response of the left M. orbicularis oculi using Ag/AgCl electrodes (diameter 5 mm). One centre (Mannheim) used a Nihon-Kohden (Tokyo, Japan) polygraph, while the other three centres used Varioport systems (Becker-Meditec, Karlsruhe, Germany). In Mannheim, the sample rate was set to 2000 Hz and data were stored in intervals of 15 s duration, triggered by the stimulation software. At the other three centres, the sample rate was set to 1024 Hz, data were recorded continuously, and the trigger coding the onsets of trials, pictures, and startles were stored in the Varioport marker channel. In addition to the EMG, heart rate and electrodermal activity were measured but not reported in this paper.

Startle measure and analysis: Data were analysed using both Vision Analyzer software (v1.05; Brain Products, Gilching, Germany) as well as self-written software (in Borland Delphi). EMG data were resampled to 1000 Hz and filtered using a 30-500 Hz bandpass. Data quality was visually controlled by an experienced electrophysiologist in a time window from -3 to +3 s around startle onset and datasets showing excessive artefacts were excluded. Then the EMG was segmented from −200 ms to 500 ms around the startle onset and rectified. Only segments that showed a clearly visible startle reflex in the range of 20 ms to 200 ms after stimulus onset were included in further analysis. Segments showing noise and/or EMG bursts outside this time range – unless the amplitude of these bursts and noise was considerably smaller (< 50%) than that of the startle EMG – as well as signals outside the amplifier range, were discarded. Data of subjects with fewer than four trials remaining in one stimulus category were excluded from analysis. The percentage of included trials did not differ ($p > 0.7$) between stimulus categories (including first and second measurement; range 87% to 88.9%). As an estimation of the signal-to-noise ratio, the mean EMG activity outside the startle time window (-200 to -20 and 180 to 500 ms) was computed for the rectified data and set in relation to the startle amplitude. This EMG activity, given in percentage of the startle amplitude, was 11.6% (alcohol pictures), 10.9% (negative), 10.4% (neutral), 10.8% (positive), 11.1% (no picture cue). Trials were averaged separately for each stimulus category and smoothed using a 10-point moving average. After baseline correction (−120 to −20 ms) to remove the offset shift resulting from the rectifying procedure, the peak amplitude was determined within the time window 20 to 180 ms.

In order to explore whether patients respond differently to alcohol-associated cues – i.e. either more appetitively or more aversively – their alcohol-cue startle response was compared with the other stimulus categories (alc minus negative, alc minus neutral, and alc minus positive). Our results revealed that our sample of patients did not exhibit the uniform valence-modulated startle response (positive < neutral < negative) that is usually found in healthy subjects. Therefore, to obtain a more robust (and conservative) reference value by reducing the variance between picture categories, the percentage of alcohol-related cues relative to the mean of negative, neutral, and positive cues was computed:

$NNP_{(mean)}=(negative+neutral+positive)/3;$
$ALC\text{-}NNP[\%]=200\times(alcohol-NNP)/(alcohol+NNP).$

Negative values indicate that patients perceive alcohol stimuli as appetitive and positive values indicate aversive reactions.

As is often the case in complex psychophysiological experiments on acutely ill patients, many artefacts were found. Also affecting the quality of the data was the fact that we were not able to conduct the experiment in a strictly controlled laboratory setting, but instead had to conduct it under outpatient conditions in a multicenter environment. As a consequence, out of 298 patients of whom we collected baseline data, only 74 yielded data of sufficient quality to be entered into the final analyses. This high dropout rate was due to two main reasons: firstly, data from 135 patients had to be excluded either because physiological or non-physiological noise was too high, a startle response was missing, or files were corrupted. Secondly, we had to exclude, as a precautionary measure, a subset of our data (n=89) which was recorded using a specific configuration (data stored in epochs, in combination with the first version of the stimulation software which generated individually randomized stimulation sequences). Due to technical problems, this data showed partly occurring synchronization errors between stimulation and physiological data that we were unable to correct for reliably. This problem was not an issue for the measurements taken using the Varioport system, which recorded EMG and stimulus markers continuously, thereby enabling us to correct for any possible synchronisation errors offline. There were 33 patients for whom we were able to analyse startle data from both measurements.

Statistical Methods

We analysed differences in picture ratings and startle responses between the stimulus categories using one-way repeated-measures ANOVAs as well as the Newman-Keuls test

for pairwise post-hoc comparisons, as appropriate. To investigate correlations between psychometric data and relative startle amplitude on alcohol pictures, we computed Pearson's *r*. Subsequent steps of the analyses aimed to identify startle related differential medication effects on relapse risk and abstinence during the 90 day medication phase. We therefore ran a Cox regression, including the factor medication and the continuous covariate relative startle response (ALC-NNP) as main effects as well as an interaction term between the two, with time to first heavy-drinking day as the failure time outcome variable, as well as a logistic regression using the abstinence status (relapse to heavy drinking: yes/no) as a binary outcome. The analyses that included drinking outcomes referred to the observation period of 3 months.

2.2.4 Results

A comparison of the subsample used in this study with the complete sample of the PREDICT study (N=426; Mann et al., 2013) revealed no significant differences in demographic or psychometric data (variable see Table 1), ensuring that dropouts due to artefactual startle data were random and did not bias the sample.

Rating of stimulus material

Picture rating scores were analysed using separate ANOVAs for valence, arousal, and craving, with picture category as the repeated-measures factor. In order to ensure better comparison between picture-related startle responses and picture ratings of each category, only those eight pictures from each category that had been presented together with a startle probe were included in the analyses. The rating results comprising all 12 pictures showed only marginal differences. For valence, the main effect was the following: $F(3, 219)=62.6$, $p<0.001$, $eta^2=0.46$. Pairwise post-hoc comparisons revealed a significantly more-aversive valence for negative pictures, compared to all other conditions, and a significantly more-appetitive valence for positive stimuli than for alcohol picture stimuli (Table 2). For arousal, the main effect was the following: $F(3, 219)=59.1$, $p<0.001$, $eta^2=0.45$. Arousal values were lowest for alcohol pictures, followed by neutral, positive, and negative pictures. Arousal values for all categories were significantly different from each other. Craving scores were generally very low (about 5 on a scale from 0–100) without any statistically significant differences. In summary, alcohol-related pictures were scored as “more neutral” than the

neutral pictures (valence closest to zero and lowest self rated arousal). An additional item analysis (each picture analysed separately) found no outliers in the stimulus material.

Table 2

Patients' ratings (mean and SD) of the four stimulus categories for valence, arousal, and craving. The rating scales (visual analogue scale) ranged from 0 (lowest) to 100 (highest score) for arousal and craving. Valence ratings are presented from −50 (most negative) to +50 (most positive). For valence, negative stimuli were significantly different from all other categories. Furthermore, positive and alcohol stimuli differed significantly. For arousal, all pairwise comparisons were significantly different. For craving, no statistical differences were found.

	Alcohol	**Negative**	**Neutral**	**Positive**
Valence	4.8 ±23	−26.6 ±20.1	9.9 ±16.6	15.5 ±22.6
Arousal	28.7 ±19.9	63.3 ±25	40.6 ±17.7	48.7 ±19.2
Craving	5.8 ±9.4	4.2 ±6.5	5.7 ±9.8	6.8 ±12.9

Physiological data

At baseline, group-mean startle responses (peak amplitudes, *N*=74) were the lowest for the alcohol-related cues and for the startle trials without picture stimuli. Negative and neutral cues revealed the largest startle-response amplitudes and positive cues were in between (Figure 1a, 1b). An ANOVA with stimulus category as the repeated-measures factor revealed a significant main effect ($F(4, 292)=4.83$, $p<0.001$, $eta^2=0.062$). Post-hoc comparisons revealed significantly smaller startle amplitudes for alcohol cues than for negative, neutral, and positive cues. In addition, the amplitudes of un-cued startles were significantly smaller than those for the negative and the neutral conditions.

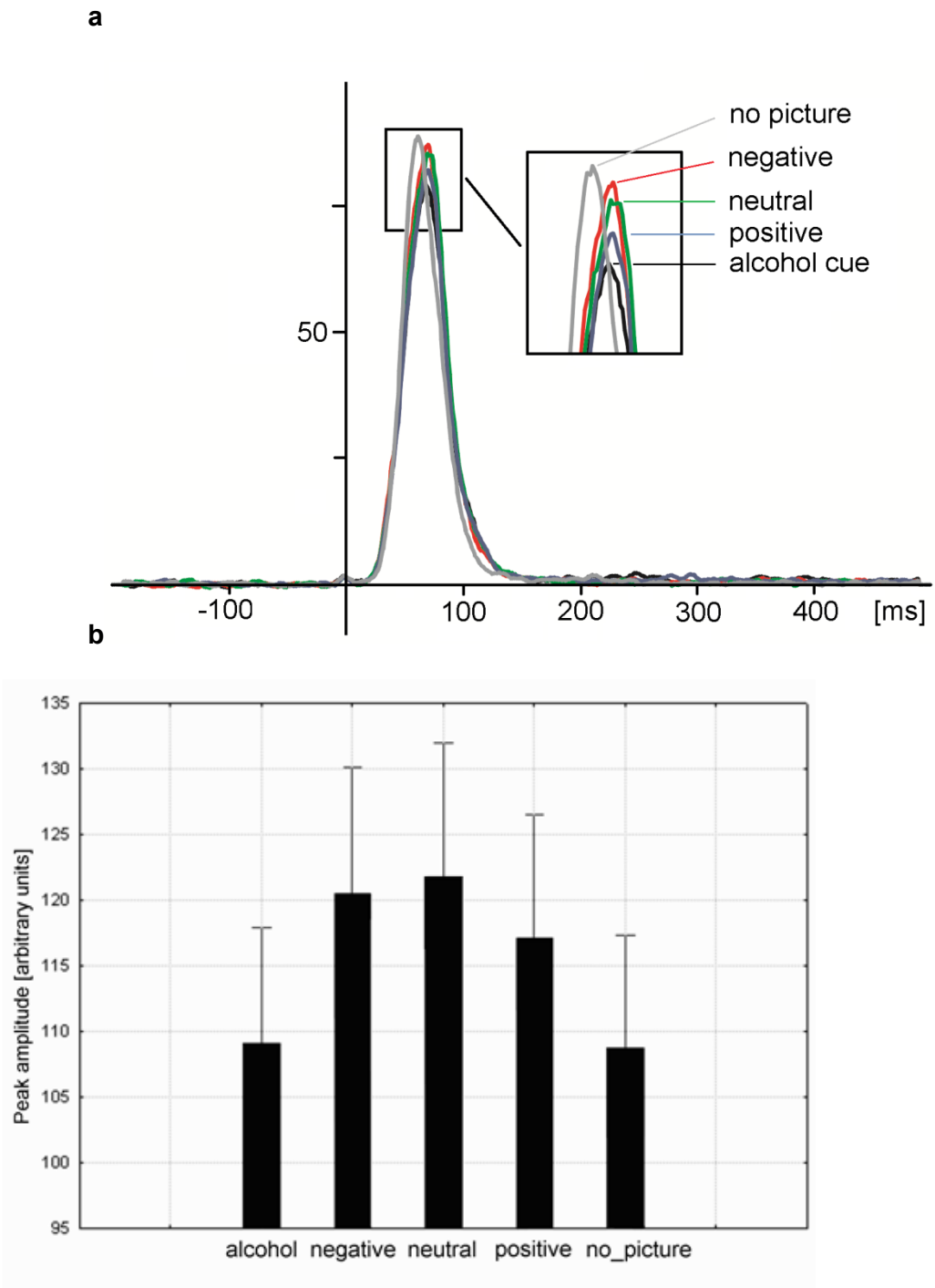


Figure 1. a: Startle responses averaged across subjects (grandaverage) for the different stimulus categories. The individual averages were smoothed with a 10-point moving average beforehand. Stimulus onset (startle probe) at time zero (indicated by the vertical line). The peak amplitudes are smaller in the grandaverages than the peak amplitudes measured in the individual averages which entered statistical analysis (Fig. 1b) for methodological reasons. The averaging procedure attenuates the amplitudes of the grandaverages due to the

individual differences in peak latencies (latency jitter). In contrast to Fig. 1b the amplitude of the uncued startle is larger because the variance in the peak latencies was smaller in that condition compared to the others. Mean and Standard Deviations of the peak latencies were 50.5 ms +- 11.4 (alcohol cues), 50.1 +-12.5 (negative), 50.4 +-13.2 (neutral), 51.3 +- 13.3 (positive), 46.9 +-10.9 (no picture). b: Startle responses (peak amplitudes and standard error) for the different stimulus categories at baseline (*N*=74). Significantly smaller responses were found for the alcohol-related stimuli and startles given without a cue image.

To compare startle responses between baseline and follow-up measure, we used an ANOVA with time and stimulus category as repeated-measures factors. Due to the high dropout rate, data for both measurement points were only available for 33 patients (Figure 2). This analysis did not reveal a significant main effect for time ($F(1, 32)=2.17$; $p=0.15$), neither an interaction effect between time and stimulus category ($F(4, 128)=0.41$, $p=0.80$). The main effect for stimulus category was significant ($F(4, 128)=9.74$, $p<0.001$, $eta^2=0.23$). The post-hoc comparisons replicated the pattern of the baseline measure. In addition, the amplitude of the positive condition was significantly smaller than those of the neutral condition.

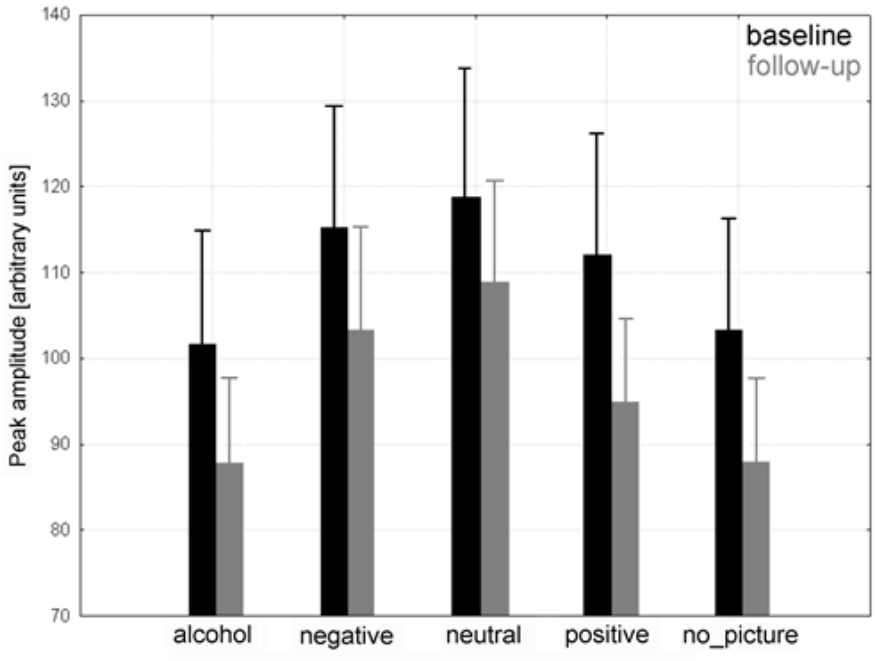


Figure 2. Startle responses (peak amplitudes and standard error) for the different stimulus categories and both measures (baseline, follow-up: 21 days later). Although the sample size at follow p was smaller (*n*=33) compared to baseline, the modulation of the startle response by stimulus category was similar. The visible decrease of startle amplitudes from baseline to follow-up was not significant.

Our analysis of the correlations between patients' relative startle responses (ALC-NNP) and both their picture ratings as well as their OCDS scores revealed no significant results.

The percentage of abstinent days in the 90-day period prior to the last drinking day (Form 90) was significantly correlated to ALC-NNP (r=0.32, p=0.005; N=74, Figure 3). This finding indicates a high number of abstinent days prior to the last drinking day to be related to aversive responses to alcohol-related stimuli.

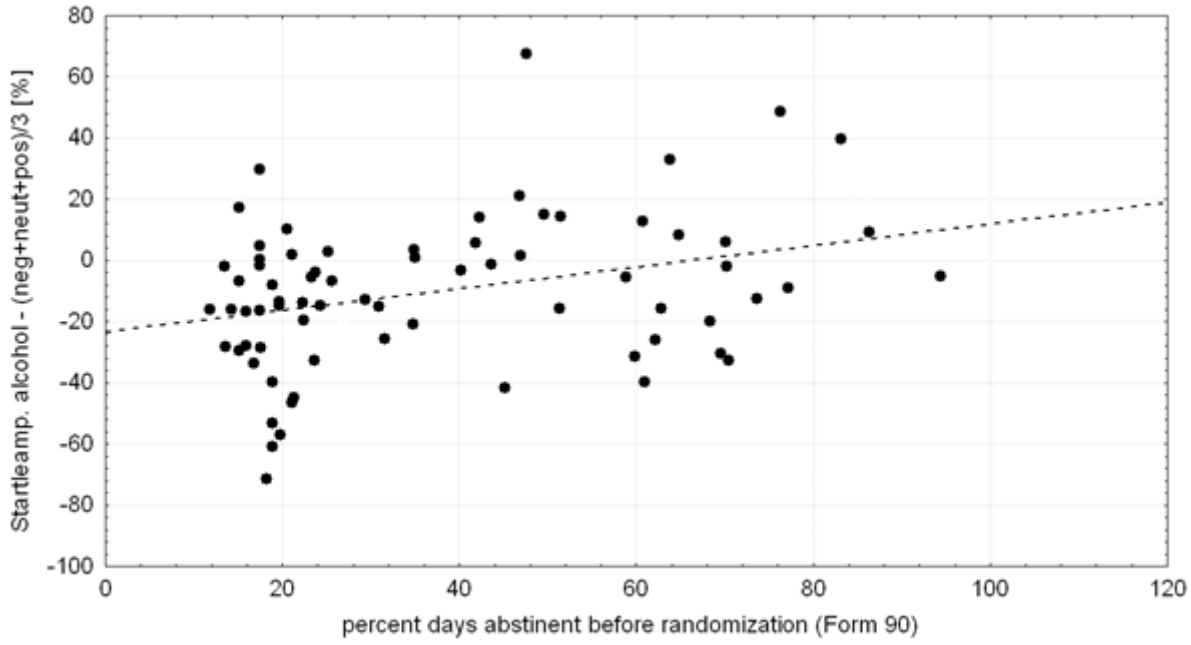


Figure 3. The percentage of abstinent days prior to the last drinking day (FORM90) is significantly correlated to relative startle response (ALC-NNP; r=0.32, p=0.005). A higher percentage of abstinent days is related to a more aversive response to alcohol-related stimuli.

In order to investigate the effects of pharmacological treatment, we ran a Cox regression with two main effects (ALC-NNP as a continuous covariate and medication as a factor) together with an interaction effect between the two, for time to first heavy drinking day as the dependent variable. This regression revealed a significant interaction (χ^2=6.64, df=2, p=0.036), pointing to a differential medication effect (Table 3).

Table 3

Results of a Cox regression for time to first heavy drinking day

Effect	**χ2**	**DF**	**p-value**
Relative startle response ALC-NNP	0.35	1	0.554
Medication	1.96	2	0.375
Medication*ALC-NNP	6.64	2	0.036

The more aversive a patient's startle response, the lower the hazard ratio of acamprosate (ACAM) vs. naltrexone (NALT). Acamprosate thus seems to be more effective than naltrexone for those patients who respond aversively to alcohol cues. On the other hand, those who respond more appetitively to alcohol-related stimuli seem to be better served by naltrexone. Comparing the outcomes of the two medication groups to those of the placebo group did not show any significant differences, perhaps owing in part to the small size of 12 patients in the placebo group. Figure 4 illustrates this significant interaction, showing the estimated log hazard-ratio curves for relative startle response by treatment arm. Inspection of the hazard ratios for ACAM vs. NALT and their 95% confidence intervals (CI) at different values of the covariate ALC-NNP revealed two different things. On the one hand, at covariate values lower than or equal to −42 (appetitive region) the hazard ratio was bigger than 1, while the CI did not include 1. This indicated that NALT treatment had an advantage in this extremely appetitive end of the startle-response scale. On the other hand, at values greater than or equal to 18, ACAM vs. NALT hazard ratios were smaller than 1, while the CI again did not include 1, suggesting ACAM to have an advantage in this extremely aversive end of the startle-response scale. The dashed vertical lines in Figure 4 mark these limits. The area in between can be interpreted as a more neutral region. The distance between the ACAM and the NALT curve at the outer ends of the startle response scale sheds light on the size of the hazard ratio ACAM vs. NALT. At the negative (appetitive) end, a reduced risk for NALT is indicated by an increasing distance, while at the positive (aversive) end the advantage of ACAM is more pronounced, with a larger distance between the curves. Considering the interaction effect from another perspective, by focusing on the differential effects of relative

startle response (ALC-NNP) in the three medication groups, we found the hazard ratio for ALC-NNP to be significant only in the acamprosate group, where a unit increase in the ALC-NNP was associated with a 3.8 percentage-point decrease in the hazard ratio. It should be emphasized that all these calculations were based on a Cox model with two main effects and an interaction term.

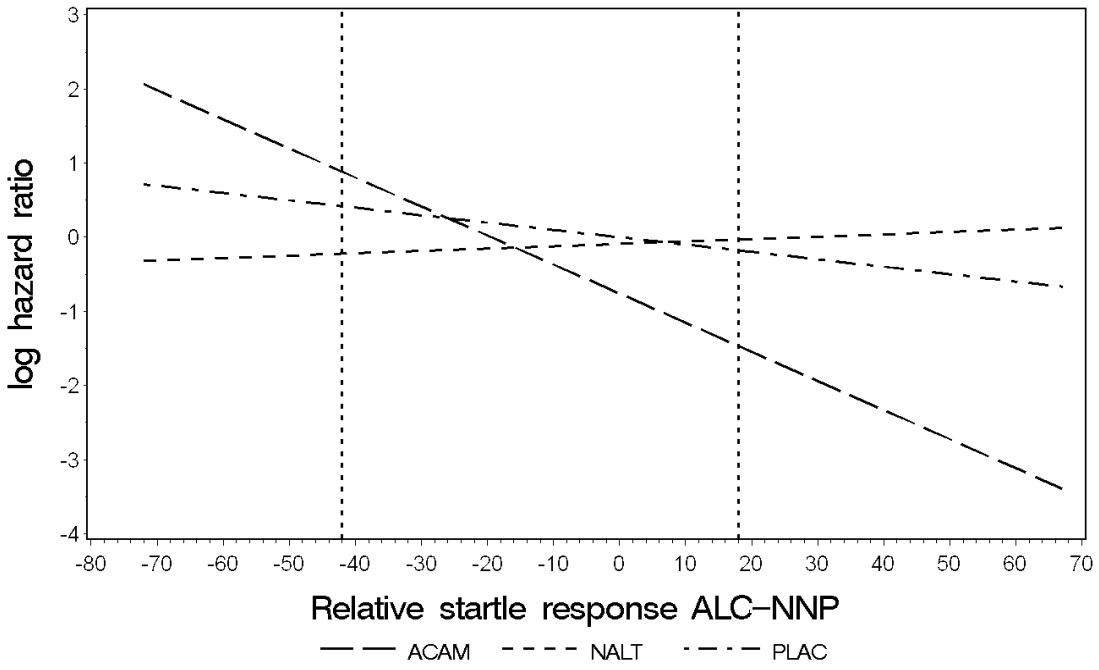


Figure 4. Estimated log hazard ratio curves for the relative startle response (ALC-NNP) by medication (ACAM *n*=27, NALT *n*=35, PLAC *n*=12) as a linearized measure of relapse risk. The distance between the ACAM and the NALT curve at the outer ends of the startle response scale gives information about the size of the hazard ratio ACAM vs. NALT. At the negative (appetitive) end a reduced risk for NALT is shown by an increasing distance while at the positive (aversive) end the advantage of ACAM is more pronounced with a larger distance between the curves.

We found similar differential effects of ALC-NNP with a second statistical approach using a logistic regression with abstinence status 90 days after randomization (abstinent: yes or no) as the binary dependent variable. In a model similar to the one described above (two main effects and an interaction term), we found a tendency for the interaction (χ^2=5.11, d.f.=2, *p*=0.078), again indicating acamprosate to be preferable to naltrexone for more aversive responders. Inspection of the confidence intervals for the ACAM vs. NALT odds ratios at

different values of the covariate yielded results similar to those found using the Cox model: At startle-response values greater than or equal to 10, the estimated ACAM vs. NALT odds ratio was smaller than 1, while the CI, as before, did not include 1, indicating that acamprosate may be preferable in the region of extremely aversive startle responses. Again, according to the results of the Cox regression, a more aversive startle-response seemed to reduce the risk in the ACAM group.

2.2.5 Discussion

In the present study the affective modulated eyeblink startle response was used as an implicit (not voluntary controlled) measure to investigate how alcohol-dependent patients respond emotionally on a psychophysiological level to alcohol-related stimuli. Our sample involved alcohol dependent inpatients, who had been receiving medical detoxification with a number of psychotherapeutic interventions for a minimum of two weeks. The individual- and group counselling might have helped them to feel more secure, comfortable and understood. This might be an explanation for the normal range of BDI scores of our sample.

The startle amplitudes of the alcohol-dependent patients were significantly smaller in response to alcohol-related cues than to negative, neutral, and positive pictures. These results are comparable to those of previous studies, which also found smaller startle amplitudes on alcohol cues compared to negative stimuli (Heinz et al., 2003), or compared to negative and neutral stimuli (Gruesser et al., 2002; Loeber et al., 2007; Mucha et al., 2000; Rubio et al., 2013). In contrast to our findings, these studies did not report significant differences between positive and alcohol-related stimuli. Furthermore, we only found small, non-significant differences between negative, neutral, and positive stimuli (which is partly confirmed by the subjective valence ratings of the stimuli, showing no significant difference between neutral and positive stimuli). Similar results in alcohol dependent patients have also been observed by Gruesser et al. (2002) and Rubio et al. (2013), neither of whom found any significant differences between these three standard categories' startle responses. This is in contrast to the validated "classical" emotionally modulated startle-response pattern (negative > neutral > positive) (Lang et al., 1990). According to this, only Mucha et al. (2000) and Loeber et al. (2007) described significant main effects between categories in alcohol dependent patients, without explicitly referring to differences between the non-alcohol-related categories. One explanation might be that the patients' motivational attention is elevated for alcohol stimuli, which might distract their attention away from the other emotional stimuli (Bradley et al., 2006). Bradley et al. (2006) postulate that attention to aversive stimuli

activates the defensive system, and leads to increased sensory input. As the stimulus becomes increasingly aversive, the defensive system mobilises available resources to elevate defensive activation. This process increases the startle reflex on aversive stimuli. In contrast, appetitive stimuli may act as inhibitors of the defensive system, thus decreasing startle amplitudes (Leite et al., 2012). Hence, alcohol triggers may distract motivational attention away from affective stimuli, reducing aversive psychophysiological reactions and neutralizing appetitive ones for alcohol dependent patients. Consequently, these processes lead to an amplitude adjustment of the reactions between the positive and negative stimulus categories.

Our findings further indicate a high number of abstinent days prior to the last drinking day to be related to aversive responses to alcohol-related stimuli. Regarding Figure 3, the distribution of the percent days abstinent revealed a surprisingly high percentage of patients (about 24 percent) with more than 60 percent abstinent days. One explanation might be that this high number might be influenced by previously received inpatient detoxification treatment. 64.3 percent of the total PREDICT sample reported to suffer from alcohol withdrawal symptoms (Mann et al., 2013), what might increase the need for detoxification treatment. The sample of the present study can be considered as representative in this respect, because the abstinent days before the last drinking day did not differ from the main PREDICT sample (see also table 1).The pattern of startle reactions that we found at baseline was repeated at the second measurement point. This finding confirmed Loeber et al.'s (2007) result that patients' startle responses are stable over time, indicating that this psycho-physiological cue-reactivity paradigm measures a trait rather than a state characteristic of cue-related emotional valence.

In contrast to the relative startle response (ALC-NNP) to alcohol stimuli, self-ratings of the valence, arousal and craving induced by alcohol-related pictures could be interpreted that these pictures were scored as even "more neutral" than the neutral pictures. Furthermore, in line with the findings of Grüsser et al. (2002), we did not find a significant correlation between alcohol-related startle response and either picture ratings or general craving, as measured by the OCDS, during the previous seven days. Also, other cue-reactivity paradigms (including some using fMRI) have not reported any association between cue-induced neural reactions and subjective craving (Hermann et al., 2006; Tapert et al., 2004; Vollstädt-Klein et al., 2011). These results show that physiological measures, which can hardly be influenced by the subject, are more sensitive in the emotional evaluation of alcohol cues. It could be also argued that these findings may be associated with a tendency toward self-beneficial

description. Due to the clinical setting, the patients tend to respond in a socially desirable way (Dillman, 2006; Kimberlin & Winterstein, 2008).

Although we did not find a significant medication effect of acamprosate or naltrexone over placebo in the overall sample of the PREDICT study (Mann et al., 2013), we investigated the hypotheses of differential medication effects based on different psychophysiological cue-reactivity according to the study protocol (Mann et al., 2009). The results of our analysis of the differing effects of the two medications indicate that patients with aversive startle responses to alcohol-related cues are likely to have lower relapse risk when treated with acamprosate than when they are treated with naltrexone. Although the literature reveals that acamprosate acts as a direct functional glutamate antagonist, there is one current study challenging these findings by proposing that acamprosate modulates the glutamatergic system via calcium (Spanagel et al., 2014). However, the exact mechanisms of acamprosate are still under debate (Heilig, 2014).

Furthermore, our results indicate that the more appetitive patients' startle responses were, the more beneficial it was to treat them with naltrexone over acamprosate. All these findings support our initial hypotheses but it has to be noted that they appeared only at the extremes of the scale for alcohol-cue-related startle responses. Nevertheless, the result indicating that the two medications have different applicability, may help categorise patients into subgroups for differential treatment options, although future randomized controlled studies will need to confirm that finding. Our findings are of special interest in the wake of the many endorsements of a personalized-medicine approach to treating alcoholism in the recent literature (Oslin et al., 2003; Heilig et al., 2011; Mann and Hermann, 2010).

Limitations of this study include the relatively small sample sizes of the medication subgroups as well as the fact that the outpatient treatment setting resulted in a high number of cases being lost due to artefacts. However, the facts that (i) demographic and psychometric data between our subsample and the complete sample of the PREDICT study (Mann et al., 2013) were not (significantly) different, and (ii) that the startle response patterns at baseline were similar for all 74 subjects and the subset of 33 subjects who participated at both measurements, indicate that dropouts due to artefactual startle data were random and did not bias the sample. Furthermore, although the startle response effects we found were significant, their effect sizes were small. This indicates that (alcohol-related) picture stimuli only induce moderate cue-reactivity effects compared to smelling or even consuming the real substances.

Conclusion

The results of the present study indicate that the modulated eyeblink startle response enables to discover differences in emotional processing of alcohol related cues in alcohol dependent patients. These psychophysiological inter-individual differences are also associated with a differential medication effect which might encourage further studies to use biological characteristics to stratify patients as a step towards individualized treatment for alcohol dependence. Furthermore, our findings of non-significant associations between alcohol-related startle response and either picture ratings or general craving may suggest that the psycho-physiological measure includes unconscious processing, which is more sensitive in the emotional evaluation of alcohol cues compared to conscious self-ratings, which additionally can be distracted by biased decisions.

Acknowledgements

The PREDICT study was supported by the Federal Ministry of Education and Research (BMBF, 01EB0110, 01EB0111, 01EB0112) and the German federal state of Baden Württemberg. The PREDICT study team consisted of more than 60 team members working at 4 different sites. Special thanks go to the following people who were involved in data collection:

University of Freiburg, Dept. of Psychiatry & Psychotherapy: Michael Berner, MD; Rigo Brück, PhD; Katrin Frick, PhD; Martin Härter, MD; Barbara Loessl, MD;

University of Tübingen, University Hospital for Psychiatry & Psychotherapy: Anil Batra, MD; Monika Haug, MD; Hubertus Magnus Friederich, MD; Robert Smolka, MD.

3 GENERAL DISCUSSION

3.1 Overall discussion and outlook

The two studies could largely answer the scientific questions posed at the beginning. Three brain regions (VS, OFC and vACC) could be identified to be associated with relapse in an exploratory pre-analysis by classical SPM correlations. This is in line with Schacht et al. (2013) who found two brain regions (VS and OFC) to be involved in reward valuation, and with Beck et al. (2012) who showed brain structure and function in these regions to be associated with relapse. In the following survival analyses of the aggregated fMRI measures several ROI data-aggregation measures were found to have significant effects on time to first severe relapse. The hypothesis on cue reactivity in the mesolimbic reward system (1a) was confirmed by finding that higher alcohol associated cue reactivity is related to an increased relapse risk in patients suffering from AUD. The fMRI cue reactivity in the VS proved to be suitable as prognostic factor for relapse. The corresponding hypothesis on the suitability of aggregated fMRI measures (1b) could be verified by identifying a data-aggregation measure most appropriate as a relapse predictor which is combining the spatial extent of cue-induced brain activation with the intensity of this activation. This measure showed the largest differences in predictive capacity for relapse in the VS in terms of explained variation. Unlike classical linear regression models there is no unique measure of predictive accuracy in Cox regression. Here the measure “proportion of explained variation” (PEV) (Heinze & Schemper, 2003) is used which is in general very low (10-35%) and cannot be compared to the classical R^2.

The hypothesis on differential efficacy of the substances naltrexone and acamprosate under various startle response patterns (2) could be supported. The psychophysiological inter-individual differences are associated with a differential medication effect. AUD patients with aversive startle responses to alcohol-related cues are supposed to have lower relapse risk when treated with acamprosate than when treated with naltrexone. On the other hand results point to a more beneficial treatment with naltrexone over acamprosate, the more appetitive their startle responses appear. These conclusions reflect the findings of Mann et al. (2014) on a neurobiological level, where we showed an interaction effect between treatment and cue-induced VS activation on time to first severe relapse for the first time. We were able to confirm the idea that the efficacy of the pharmacotherapy could be increased if it was possible to target medications to subgroups of patients displaying a neurobiologically defined profile that was associated with response to a specific medication (Litten et al., 2012). The

results underline the clinical relevance of the studies conducted. Thus a stratification of patients to more homogeneous subgroups on various physiological profiles could support the idea of an individualized treatment. Thereby future studies should consider both neurobiological and psychophysiological individual patterns as moderating effects in evaluating medication effects on relapse behavior. Moreover, inconsistent findings on the influence of genetic variations (Bach et al., 2020; Schacht et al., 2017) in analyses of treatment effects suggest to incorporate genotypes in complex survival models as well as smoking status which was found to moderate the effect of medication on drinking (Schacht et al., 2017).

Additional research in the field of the assessment of cue reactivity and its role as a relapse predictor in patients suffering from AUD is essential to construct a multifactor model for understanding the neural basis of the processes associated with relapse mechanisms. Moreover, there would be progress in fMRI research in the clinical setting if the employed methods of data aggregation on the one hand and the complex survival analyses on the other hand were implemented in standard whole-brain fMRI software.

3.2 Limitations

The dropout rate in the first study was very high (42%) and included both fMRI dropouts and relapse cases. Thus the findings of this study may only be valid for the subgroup of patients who do not relapse directly after inpatient treatment. The results should be interpreted with caution and need to be replicated in future studies involving a greater sample size.

There are also some limitations with the second study. These include the relatively small sample sizes of the medication subgroups and dropouts due to artefactual startle data, although we were able to show that these dropouts were random and did not bias the sample. Furthermore, effect sizes were relatively small.

4 SUMMARY

The overall objective of the two studies was to clarify the extent to which functional neurobiological (fMRI) and psychophysiological (startle reflex) measures of cue reactivity are suitable predictors of severe relapse in individuals with AUD. This could be achieved by complex modern methods of survival analysis. It could be shown that fMRI cue reactivity in the ventral striatum (VS) is suitable as a prognostic factor for relapse in patients suffering from AUD by calculating different fMRI-based aggregation measures which are not yet implemented in standard whole-brain fMRI software. For the VS a measure combining the spatial extent of cue-induced brain activation with the intensity of this activation was found to be most appropriate as a biomarker for relapse prediction.

Furthermore the startle response as a psychophysiological measure showed a moderator effect when evaluating differential medication effects of naltrexone and acamprosate. The results suggest that AUD patients benefit more from naltrexone than from acamprosate treatment, the more appetitive their startle response appears. In contrast the findings point to lower relapse risk for patients with an aversive startle response pattern when treated with acamprosate compared to naltrexone. The findings support the idea of an individualized treatment based on differential pharmacological treatment due to different underlying biological mechanisms which can be identified by the affective modulation of the startle response.

The presented methods offer a potential for future analyses of high clinical relevance, also in areas besides addiction (i.e. psychiatry, oncology).

5 REFERENCES

Allison, P. D. (2010). *Survival Analysis Using SAS®: A Practical Guide, Second Edition. Cary, NC: SAS Institute Inc.*

Alvarez Moreno, E., Jimenez de la Peña, M., & Cano Alonso, R. (2012). Role of New Functional MRI Techniques in the Diagnosis, Staging, and Followup of Gynecological Cancer: Comparison with PET-CT. *Radiol Res Pract, 2012*, 219546. https://doi.org/10.1155/2012/219546

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders, DSM-IV-TR®Fourth edition. American Psychiatric Association, (943 pp.)*.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th ed.).* https://doi.org/10.1176/appi.books.9780890425596

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. *Alcohol Clin Exp Res, 19*(1), 92-99. https://doi.org/10.1111/j.1530-0277.1995.tb01475.x

Atzendorf, J., Rauschert, C., Seitz, N.-N., Lochbuehler, K., & Kraus, L. (2019). The Use of Alcohol, Tobacco, Illegal Drugs and Medicines An Estimate of Consumption and Substance-Related Disorders in Germany. *Deutsches Ärzteblatt International, 116*, 577–584.

Bach, P., Weil, G., Pompili, E., Hoffmann, S., Hermann, D., Vollstädt-Klein, S., Mann, K., Perez-Ramirez, U., Moratal, D., Canals, S., Dursun, S. M., Greenshaw, A. J., Kirsch, P., Kiefer, F., & Sommer, W. H. (2020). Incubation of neural alcohol cue reactivity after withdrawal and its blockade by naltrexone. *Addict Biol, 25*(1), e12717. https://doi.org/10.1111/adb.12717

Beck, A., Wüstenberg, T., Genauck, A., Wrase, J., Schlagenhauf, F., Smolka, M. N., Mann, K., & Heinz, A. (2012). Effect of brain structure, brain function, and brain connectivity on relapse in alcohol-dependent patients. *Arch Gen Psychiatry, 69*(8), 842-852. https://doi.org/10.1001/archgenpsychiatry.2011.2026

Bhattacharyya, P., Mathew, B., & Lowe M. (2013). BOLD signal change or activation volume – which one is more robust measure of fMRI activation? In: Presented at the 19th international conference on functional mapping of the human brain.

Bøvelstad, H. M., & Borgan, O. (2011). Assessment of evaluation criteria for survival prediction from genomic data. *Biom J, 53*(2), 202-216. https://doi.org/10.1002/bimj.201000048

Bradley, M. M., Codispoti, M., & Lang, P. J. (2006). A multi-process account of startle modulation during affective perception. *Psychophysiology, 43*(5), 486-497. https://doi.org/10.1111/j.1469-8986.2006.00412.x

Braus, D. F., Wrase, J., Grüsser, S., Hermann, D., Ruf, M., Flor, H., Mann, K., & Heinz, A. (2001). Alcohol-associated stimuli activate the ventral striatum in abstinent alcoholics. *J Neural Transm (Vienna), 108*(7), 887-894. https://doi.org/10.1007/s007020170038

Brett, M., Anton, J., Valabrègue, R., & Poline, J. B. (2002). Region of interest analysis using an SPM toolbox. Neuroimage, 16, 1140-1141.

Brody, A. L., Mandelkern, M. A., Olmstead, R. E., Jou, J., Tiongson, E., Allen, V., Scheibal, D., London, E. D., Monterosso, J. R., Tiffany, S. T., Korb, A., Gan, J. J., & Cohen, M. S. (2007). Neural substrates of resisting craving during cigarette cue exposure. *Biol Psychiatry, 62*(6), 642-651. https://doi.org/10.1016/j.biopsych.2006.10.026

Chick, J. (1995). Acamprosate as an aid in the treatment of alcoholism. *Alcohol Alcohol, 30*(6), 785-787.

Childress, A. R., Mozley, P. D., McElgin, W., Fitzgerald, J., Reivich, M., & O'Brien, C. P. (1999). Limbic activation during cue-induced cocaine craving. *Am J Psychiatry, 156*(1), 11-18. https://doi.org/10.1176/ajp.156.1.11

Cook, E. W., Davis, T. L., Hawk, L. W., Spence, E. L., & Gautier, C. H. (1992). Fearfulness and startle potentiation during aversive visual stimuli. *Psychophysiology, 29*(6), 633-645. https://doi.org/10.1111/j.1469-8986.1992.tb02038.x

Cox, D. R., & Oakes, D. (1984). Analysis of Survival Data (1st ed.). CRC Press. https://doi.org/10.1201/9781315137438

Dawson, D. A., Goldstein, R. B., & Grant, B. F. (2013). Differences in the profiles of DSM-IV and DSM-5 alcohol use disorders: implications for clinicians. *Alcohol Clin Exp Res, 37 Suppl 1*(0 1), E305-313. https://doi.org/10.1111/j.1530-0277.2012.01930.x

Diekhof, E. K., Kaps, L., Falkai, P., & Gruber, O. (2012). The role of the human ventral striatum and the medial orbitofrontal cortex in the representation of reward magnitude - an activation likelihood estimation meta-analysis of neuroimaging studies of passive reward expectancy and outcome processing. *Neuropsychologia, 50*(7), 1252-1266. https://doi.org/10.1016/j.neuropsychologia.2012.02.007

Dillman, D. A. (2006). Why choice of survey mode makes a difference. *Public Health Rep, 121*(1), 11-13. https://doi.org/10.1177/003335490612100106

Drummond, D. C. (2001). Theories of drug craving, ancient and modern. *Addiction, 96*(1), 33-46. https://doi.org/10.1046/j.1360-0443.2001.961333.x

Drummond, D. C., Cooper, T., & Glautier, S. P. (1990). Conditioned learning in alcohol dependence: implications for cue exposure treatment. *Br J Addict, 85*(6), 725-743. https://doi.org/10.1111/j.1360-0443.1990.tb01685.x

Drummond, D. C., Tiffany, S. T., Glautier, S. P., & Remington, B. (1995). Cue exposure in under-standing and treating addictive behaviours. In: D. C. Drummond, S. T. Tiffany, S. P. Glautier, B. Remington, B. (Eds.), Addictive Behaviour: Cue Exposure Theory and Practice. JohnWiley, Chichester, pp. 1–17.

Farb, N. A., Anderson, A. K., Bloch, R. T., & Segal, Z. V. (2011). Mood-linked responses in medial prefrontal cortex predict relapse in patients with recurrent unipolar depression. *Biol Psychiatry, 70*(4), 366-372. https://doi.org/10.1016/j.biopsych.2011.03.009

Franklin, T. R., Wang, Z., Wang, J., Sciortino, N., Harper, D., Li, Y., Ehrman, R., Kampman, K., O'Brien, C. P., Detre, J. A., & Childress, A. R. (2007). Limbic activation to cigarette smoking cues independent of nicotine withdrawal: a perfusion fMRI study.

Neuropsychopharmacology, *32*(11), 2301-2309. https://doi.org/10.1038/sj.npp.1301371

Garavan, H., Pankiewicz, J., Bloom, A., Cho, J. K., Sperry, L., Ross, T. J., Salmeron, B. J., Risinger, R., Kelley, D., & Stein, E. A. (2000). Cue-induced cocaine craving: neuroanatomical specificity for drug users and drug stimuli. *Am J Psychiatry, 157*(11), 1789-1798. https://doi.org/10.1176/appi.ajp.157.11.1789

Geier, A., Mucha, R. F., & Pauli, P. (2000). Appetitive nature of drug cues confirmed with physiological measures in a model using pictures of smoking. *Psychopharmacology (Berl), 150*(3), 283-291. https://doi.org/10.1007/s002130000404

George, M. S., Anton, R. F., Bloomer, C., Teneback, C., Drobes, D. J., Lorberbaum, J. P., Nahas, Z., & Vincent, D. J. (2001). Activation of prefrontal cortex and anterior thalamus in alcoholic subjects on exposure to alcohol-specific cues. *Arch Gen Psychiatry, 58*(4), 345-352. https://doi.org/10.1001/archpsyc.58.4.345

Glautier, S., Drummond, C., & Remington, B. (1994). Alcohol as an unconditioned stimulus in human classical conditioning. *Psychopharmacology (Berl), 116*(3), 360-368. https://doi.org/10.1007/bf02245341

Glöckner-Rist, A., Lemenager, T., Mann, K., & the PREDICT Study Research Group. (2013). Reward and relief craving tendencies in patients with alcohol use disorders: results from the PREDICT study. *Addict Behav, 38*(2), 1532-1540. https://doi.org/10.1016/j.addbeh.2012.06.018

Groenewegen, H. J., & Trimble, M. (2007). The ventral striatum as an interface between the limbic and motor systems. *CNS Spectr, 12*(12), 887-892. https://doi.org/10.1017/s1092852900015650

Grüsser, S. M., Heinz, A., & Flor, H. (2000). Standardized stimuli to assess drug craving and drug memory in addicts. *J Neural Transm (Vienna), 107*(6), 715-720. https://doi.org/10.1007/s007020070072

Grüsser, S. M., Heinz, A., Raabe, A., Wessa, M., Podschus, J., & Flor, H. (2002). Stimulus-induced craving and startle potentiation in abstinent alcoholics and controls. *Eur Psychiatry, 17*(4), 188-193. https://doi.org/10.1016/s0924-9338(02)00666-1

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*(3), 296-302. https://doi.org/10.1007/s00213-004-1828-4

Hammersley, R. (1992). Cue exposure and learning theory. *Addict Behav, 17*(3), 297-300. https://doi.org/10.1016/0306-4603(92)90035-t

Hautzinger, M., Bailer, M., Worall, H., Keller, F. (1995). *Beck-Depressions-Inventar (BDI). Testhandbuch. 2. Auflage. Hans Huber*.

Heilig, M. (2014). Acamprosate: an alcoholism treatment that may not be what we thought. *Neuropsychopharmacology, 39*(4), 781-782. https://doi.org/10.1038/npp.2013.272

Heilig, M., Goldman, D., Berrettini, W., & O'Brien, C. P. (2011). Pharmacogenetic approaches to the treatment of alcohol addiction. *Nat Rev Neurosci, 12*(11), 670-684. https://doi.org/10.1038/nrn3110

Heinz, A., Lober, S., Georgi, A., Wrase, J., Hermann, D., Rey, E. R., Wellek, S., & Mann, K. (2003). Reward craving and withdrawal relief craving: assessment of different motivational pathways to alcohol intake. *Alcohol Alcohol, 38*(1), 35-39. https://doi.org/10.1093/alcalc/agg005

Heinze, G., & Schemper, M. (2003). Comparing the importance of prognostic factors in Cox and logistic regression using SAS. *Computer Methods and Programs in Biomedicine, 71*(2), 155-163. https://doi.org/10.1016/S0169-2607(02)00077-9

Hermann, D., Smolka, M. N., Wrase, J., Klein, S., Nikitopoulos, J., Georgi, A., Braus, D. F., Flor, H., Mann, K., & Heinz, A. (2006). Blockade of cue-induced brain activation of abstinent alcoholics by a single administration of amisulpride as measured with fMRI. *Alcohol Clin Exp Res, 30*(8), 1349-1354. https://doi.org/10.1111/j.1530-0277.2006.00174.x

Hölter, S. M., & Spanagel, R. (1999). Effects of opiate antagonist treatment on the alcohol deprivation effect in long-term ethanol-experienced rats. *Psychopharmacology (Berl), 145*(4), 360-369. https://doi.org/10.1007/s002130051069

Hommer, D. W. (1999). Functional imaging of craving. *Alcohol Res Health, 23*(3), 187-196.

Kalbfleisch, J. D., & Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*. John Wiley & Sons.

Kalinowski, A., & Humphreys, K. (2016). Governmental standard drink definitions and low-risk alcohol consumption guidelines in 37 countries. *Addiction, 111*(7), 1293-1298. https://doi.org/10.1111/add.13341

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. https://doi.org/10.2146/ajhp070364

Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis : A Self-Learning Text, Third Edition*. Springer Science+Business Media, LLC.

Kotlinska, J., & Bochenski, M. (2008). The influence of various glutamate receptors antagonists on anxiety-like effect of ethanol withdrawal in a plus-maze test in rats. *Eur J Pharmacol, 598*(1-3), 57-63. https://doi.org/10.1016/j.ejphar.2008.09.026

Krystal, J. H., Webb, E., Grillon, C., Cooney, N., Casal, L., Morgan, C. A., 3rd, Southwick, S. M., Davis, M., & Charney, D. S. (1997). Evidence of acoustic startle hyperreflexia in recently detoxified early onset male alcoholics: modulation by yohimbine and m-chlorophenylpiperazine (mCPP). *Psychopharmacology (Berl), 131*(3), 207-215. https://doi.org/10.1007/s002130050285

Kwako, L. E., Momenan, R., Litten, R. Z., Koob, G. F., & Goldman, D. (2016). Addictions Neuroclinical Assessment: A Neuroscience-Based Framework for Addictive Disorders. *Biol Psychiatry, 80*(3), 179-189. https://doi.org/10.1016/j.biopsych.2015.10.024

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychol Rev, 97*(3), 377-395.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion and motivation: measuring affective perception. *J Clin Neurophysiol, 15*(5), 397-408. https://doi.org/10.1097/00004691-199809000-00004

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). *International Affective Picture System (IAPS): Technical Manual and Affective Ratings. University of Florida, Gainsville, FL*.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. Center for Research in Psychophysiology, University of Florida, Gainesville, FL.*

Leite, J., Carvalho, S., Galdo-Alvarez, S., Alves, J., Sampaio, A., & Gonçalves, O. F. (2012). Affective picture modulation: valence, arousal, attention allocation and motivational significance. *Int J Psychophysiol, 83*(3), 375-381. https://doi.org/10.1016/j.ijpsycho.2011.12.005

Leménager, T., Hill, H., Reinhard, I., Hoffmann, S., Zimmermann, U. S., Hermann, D., Smolka, M. N., Kiefer, F., Vollstädt-Klein, S., Heinz, A., & Mann, K. (2014). Association between alcohol-cue modulated startle reactions and drinking behaviour in alcohol dependent patients - results of the PREDICT study. *Int J Psychophysiol, 94*(3), 263-271. https://doi.org/10.1016/j.ijpsycho.2014.09.009

Litten, R. Z., Egli, M., Heilig, M., Cui, C., Fertig, J. B., Ryan, M. L., Falk, D. E., Moss, H., Huebner, R., & Noronha, A. (2012). Medications development to treat alcohol dependence: a vision for the next decade. *Addict Biol, 17*(3), 513-527. https://doi.org/10.1111/j.1369-1600.2012.00454.x

Loeber, S., Croissant, B., Nakovics, H., Zimmer, A., Georgi, A., Klein, S., Diener, C., Heinz, A., Mann, K., & Flor, H. (2007). The startle reflex in alcohol-dependent patients: changes after cognitive-behavioral therapy and predictive validity for drinking behavior. A pilot study. *Psychother Psychosom, 76*(6), 385-390. https://doi.org/10.1159/000107567

Loeber, S., Vollstadt-Klein, S., von der Goltz, C., Flor, H., Mann, K., & Kiefer, F. (2009). Attentional bias in alcohol-dependent patients: the role of chronicity and executive functioning. *Addict Biol, 14*(2), 194-203. https://doi.org/10.1111/j.1369-1600.2009.00146.x

Lupien, S. J., Sasseville, M., François, N., Giguère, C. E., Boissonneault, J., Plusquellec, P., Godbout, R., Xiong, L., Potvin, S., Kouassi, E., & Lesage, A. (2017). The DSM5/RDoC debate on the future of mental health research: implication for studies on human stress and presentation of the signature bank. *Stress, 20*(1), 95-111. https://doi.org/10.1080/10253890.2017.1286324

Maas, L. C., Lukas, S. E., Kaufman, M. J., Weiss, R. D., Daniels, S. L., Rogers, V. W., Kukes, T. J., & Renshaw, P. F. (1998). Functional magnetic resonance imaging of human brain activation during cue-induced cocaine craving. *Am J Psychiatry, 155*(1), 124-126. https://doi.org/10.1176/ajp.155.1.124

Mann, K., & Ackermann, K. (2000). Die OCDS-G: Psychometrische Kennwerte der deutschen Version der Obsessive Compulsive Drinking Scale. *SUCHT, 46*(2), 90-100. https://doi.org/10.1024/suc.2000.46.2.90

Mann, K., Kiefer, F., Smolka, M., Gann, H., Wellek, S., Heinz, A., & The PREDICT Study Research Team (2009). Searching for responders to acamprosate and naltrexone in alcoholism treatment: rationale and design of the PREDICT study. *Alcohol Clin Exp Res, 33*(4), 674-683. https://doi.org/10.1111/j.1530-0277.2008.00884.x

Mann, K., Lehert, P., & Morgan, M. Y. (2004). The efficacy of acamprosate in the maintenance of abstinence in alcohol-dependent individuals: results of a meta-analysis. *Alcohol Clin Exp Res, 28*(1), 51-63. https://doi.org/10.1097/01.Alc.0000108656.81563.05

Mann, K., Lemenager, T., Hoffmann, S., Reinhard, I., Hermann, D., Batra, A., Berner, M., Wodarz, N., Heinz, A., Smolka, M. N., Zimmermann, U. S., Wellek, S., Kiefer, F., Anton, R. F., & The PREDICT Study Team (2013). Results of a double-blind, placebo-controlled pharmacotherapy trial in alcoholism conducted in Germany and comparison with the US COMBINE study. *Addict Biol, 18*(6), 937-946. https://doi.org/10.1111/adb.12012

Mann, K., Vollstädt-Klein, S., Reinhard, I., Leménager, T., Fauth-Bühler, M., Hermann, D., Hoffmann, S., Zimmermann, U. S., Kiefer, F., Heinz, A., & Smolka, M. N. (2014). Predicting naltrexone response in alcohol-dependent patients: the contribution of functional magnetic resonance imaging. *Alcohol Clin Exp Res, 38*(11), 2754-2762. https://doi.org/10.1111/acer.12546

McBride, D., Barrett, S. P., Kelly, J. T., Aw, A., & Dagher, A. (2006). Effects of expectancy and abstinence on the neural response to smoking cues in cigarette smokers: an fMRI study. *Neuropsychopharmacology, 31*(12), 2728-2738. https://doi.org/10.1038/sj.npp.1301075

Miller, W. R. (1996). Form 90: *A structured assessment interview for drinking and related behaviors: Test manual.* Project MATCH Monograph Series vol. 5. National Institute on Alcohol Abuse and Alcoholism, Rockville.

Mucha, R. F., Geier, A., Stuhlinger, M., & Mundle, G. (2000). Appetitve effects of drug cues modelled by pictures of the intake ritual: generality of cue-modulated startle examined with inpatient alcoholics. *Psychopharmacology (Berl), 151*(4), 428-432. https://doi.org/10.1007/s002130000508

Niaura, R. S., Rohsenow, D. J., Binkoff, J. A., Monti, P. M., Pedraza, M., & Abrams, D. B. (1988). Relevance of cue reactivity to understanding alcohol and smoking relapse. *J Abnorm Psychol, 97*(2), 133-152. https://doi.org/10.1037//0021-843x.97.2.133

Nielsen, F. A., & Hansen, L. K. (2002). Automatic anatomical labeling of Talairach coordinates and generation of volumes of interest via the BrainMap database. NeuroImage 2: presented at the 8th annual conference on functional mapping of the human brain.

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 U S A, 87*(24), 9868-9872. https://doi.org/10.1073/pnas.87.24.9868

Ogawa, S., Menon, R. S., Tank, D. W., Kim, S. G., Merkle, H., Ellermann, J. M., & Ugurbil, K. (1993). Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. *Biophys J, 64*(3), 803-812. https://doi.org/10.1016/s0006-3495(93)81441-3

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113. https://doi.org/10.1016/0028-3932(71)90067-4

Oslin, D. W., Berrettini, W., Kranzler, H. R., Pettinati, H., Gelernter, J., Volpicelli, J. R., & O'Brien, C. P. (2003). A functional polymorphism of the mu-opioid receptor gene is associated with naltrexone response in alcohol-dependent patients. *Neuropsychopharmacology, 28*(8), 1546-1552. https://doi.org/10.1038/sj.npp.1300219

Poldrack, R. A. (2007). Region of interest analysis for fMRI. *Soc Cogn Affect Neurosci, 2*(1), 67-70. https://doi.org/10.1093/scan/nsm006

Rassnick, S., Koob, G. F., & Geyer, M. A. (1992). Responding to acoustic startle during chronic ethanol intoxication and withdrawal. *Psychopharmacology (Berl), 106*(3), 351-358. https://doi.org/10.1007/bf02245417

Reinhard, I., Lemenager, T., Fauth-Bühler, M., Hermann, D., Hoffmann, S., Heinz, A., Kiefer, F., Smolka, M. N., Wellek, S., Mann, K., & Vollstadt-Klein, S. (2015). A comparison of region-of-interest measures for extracting whole brain data using survival analysis in alcoholism as an example. *J Neurosci Methods, 242*, 58-64. https://doi.org/10.1016/j.jneumeth.2015.01.001

Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res Brain Res Rev, 18*(3), 247-291. https://doi.org/10.1016/0165-0173(93)90013-p

Robinson, M. J., & Berridge, K. C. (2013). Instant transformation of learned repulsion into motivational "wanting". *Curr Biol, 23*(4), 282-289. https://doi.org/10.1016/j.cub.2013.01.016

Rohsenow, D. J., Niaura, R. S., Childress, A. R., Abrams, D. B., & Monti, P. M. (1990). Cue reactivity in addictive behaviors: theoretical and treatment implications. *Int J Addict, 25*(7a-8a), 957-993. https://doi.org/10.3109/10826089109071030

Rubio, G., Borrell, J., Jimenez, M., Jurado, R., Grüsser, S. M., & Heinz, A. (2013). Variables involved in the cue modulation of the startle reflex in alcohol-dependent patients. *Addict Biol, 18*(1), 170-180. https://doi.org/10.1111/j.1369-1600.2011.00371.x

Saladin, M. E., Drobes, D. J., Coffey, S. F., & Libet, J. M. (2002). The human startle reflex and alcohol cue reactivity: effects of early versus late abstinence. *Psychol Addict Behav, 16*(2), 98-105. https://doi.org/10.1037//0893-164x.16.2.98

Schacht, J. P., Anton, R. F., Randall, P. K., Li, X., Henderson, S., & Myrick, H. (2011). Stability of fMRI striatal response to alcohol cues: a hierarchical linear modeling approach. *Neuroimage, 56*(1), 61-68. https://doi.org/10.1016/j.neuroimage.2011.02.004

Schacht, J. P., Anton, R. F., & Myrick, H. (2013). Functional neuroimaging studies of alcohol cue reactivity: a quantitative meta-analysis and systematic review. *Addict Biol, 18*(1), 121-133. https://doi.org/10.1111/j.1369-1600.2012.00464.x

Schacht, J. P., Randall, P. K., Latham, P. K., Voronin, K. E., Book, S. W., Myrick, H., & Anton, R. F. (2017). Predictors of Naltrexone Response in a Randomized Trial: Reward-Related Brain Activation, OPRM1 Genotype, and Smoking Status. *Neuropsychopharmacology, 42*(13), 2640-2653. https://doi.org/10.1038/npp.2017.74

Schad, L. R. (2002a). [Functional magnetic resonance tomography (fMRI). 1: Basic principles and measuring techniques]. *Radiologe, 42*(8), 659-666; quiz 667-659. https://doi.org/10.1007/s00117-002-0788-0 (Funktionelle Magnetresonanztomographie (fMRT). Teil 1: Grundlagen und Messtechniken.)

Schad, L. R. (2002b). [Functional magnetic resonance tomography (fMRI). 2: Data analysis and applications]. *Radiologe, 42*(9), 756-763; quiz 764. https://doi.org/10.1007/s00117-002-0789-z (Funktionelle Magnetresonanztomographie (fMRT). Teil 2: Datenanalyse und Anwendungen.)

Schemper, M., & Henderson, R. (2000). Predictive Accuracy and Explained Variation in Cox Regression. *Biometrics, 56*(1), 249-255. http://www.jstor.org.ezproxy.medma.uni-heidelberg.de/stable/2677129

Schneider, F., Habel, U., Wagner, M., Franke, P., Salloum, J. B., Shah, N. J., Toni, I., Sulzbach, C., Hönig, K., Maier, W., Gaebel, W., & Zilles, K. (2001). Subcortical correlates of craving in recently abstinent alcoholic patients. *Am J Psychiatry, 158*(7), 1075-1083. https://doi.org/10.1176/appi.ajp.158.7.1075

Schwartz, T. L., Siddiqui, U. A., Raza, S., & Costello, A. (2010). Acamprosate calcium as augmentation therapy for anxiety disorders. *Ann Pharmacother, 44*(12), 1930-1932. https://doi.org/10.1345/aph.1P353

Smolka, M. N., Bühler, M., Klein, S., Zimmermann, U., Mann, K., Heinz, A., & Braus, D. F. (2006). Severity of nicotine dependence modulates cue-induced brain activity in regions involved in motor preparation and imagery. *Psychopharmacology (Berl), 184*(3-4), 577-588. https://doi.org/10.1007/s00213-005-0080-x

Spanagel, R., Vengeliene, V., Jandeleit, B., Fischer, W. N., Grindstaff, K., Zhang, X., Gallop, M. A., Krstew, E. V., Lawrence, A. J., & Kiefer, F. (2014). Acamprosate produces its anti-relapse effects via calcium. *Neuropsychopharmacology, 39*(4), 783-791. https://doi.org/10.1038/npp.2013.264

Strother, S. C. (2006). Evaluating fMRI preprocessing pipelines. *IEEE Eng Med Biol Mag, 25*(2), 27-41. https://doi.org/10.1109/memb.2006.1607667

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*(1), 33-50. https://doi.org/10.1016/j.addbeh.2003.07.003

Tiffany, S. T., Carter, B. L., & Singleton, E. G. (2000). Challenges in the manipulation, assessment and interpretation of craving relevant variables. *Addiction, 95 Suppl 2*, S177-187. https://doi.org/10.1080/09652140050111753

Tiffany, S. T., & Conklin, C. A. (2000). A cognitive processing model of alcohol craving and compulsive alcohol use. *Addiction, 95*(8 Suppl 2), 145-153. https://doi.org/10.1080/09652140050111717

Tost, H., Meyer-Lindenberg, A., Klein, S., Schmitt, A., Hohn, F., Tenckhoff, A., Ruf, M., Ende, G., Rietschel, M., Henn, F. A., & Braus, D. F. (2006). D2 antidopaminergic modulation of frontal lobe function in healthy human subjects. *Biol Psychiatry, 60*(11), 1196-1205. https://doi.org/10.1016/j.biopsych.2006.04.014

Verheul, R., van den Brink, W., & Geerlings, P. (1999). A three-pathway psychobiological model of craving for alcohol. *Alcohol Alcohol, 34*(2), 197-222. https://doi.org/10.1093/alcalc/34.2.197

Vollstädt-Klein, S., Loeber, S., Kirsch, M., Bach, P., Richter, A., Bühler, M., von der Goltz, C., Hermann, D., Mann, K., & Kiefer, F. (2011). Effects of cue-exposure treatment on neural cue reactivity in alcohol dependence: a randomized trial. *Biol Psychiatry, 69*(11), 1060-1066. https://doi.org/10.1016/j.biopsych.2010.12.016

Vollstädt-Klein, S., Wichert, S., Rabinstein, J., Buhler, M., Klein, O., Ende, G., Hermann, D., & Mann, K. (2010). Initial, habitual and compulsive alcohol use is characterized by a shift of cue processing from ventral to dorsal striatum. *Addiction, 105*(10), 1741-1749. https://doi.org/10.1111/j.1360-0443.2010.03022.x

Volpicelli, J. R., Clay, K. L., Watson, N. T., & O'Brien, C. P. (1995). Naltrexone in the treatment of alcoholism: predicting response to naltrexone. *J Clin Psychiatry, 56 Suppl 7*, 39-44.

Vrana, S. R., Spence, E. L., & Lang, P. J. (1988). The startle probe response: a new measure of emotion? *J Abnorm Psychol, 97*(4), 487-491. https://doi.org/10.1037//0021-843x.97.4.487

Weiss, F., Ciccocioppo, R., Parsons, L. H., Katner, S., Liu, X., Zorrilla, E. P., Valdez, G. R., Ben-Shahar, O., Angeletti, S., & Richter, R. R. (2001). Compulsive drug-seeking behavior and relapse. Neuroadaptation, stress, and conditioning factors. *Ann N Y Acad Sci, 937*, 1-26. https://doi.org/10.1111/j.1749-6632.2001.tb03556.x

Wittchen, H.-U., Zaudig, M., & Fydrich, T. (1997). *SKID. Strukturiertes Klinisches Interview für DSM-IV. Achse I und II. Handanweisung*. Hogrefe. http://hdl.handle.net/11858/00-001M-0000-000E-AAB2-2

Wojtowicz, M., Mazerolle, E. L., Bhan, V., & Fisk, J. D. (2014). Altered functional connectivity and performance variability in relapsing-remitting multiple sclerosis. *Mult Scler, 20*(11), 1453-1463. https://doi.org/10.1177/1352458514524997

World Health Organization (1992). *The ICD-10 classification of mental and behavioural disorders : clinical descriptions and diagnostic guidelines*. World Health Organization. https://apps.who.int/iris/handle/10665/37958

World Health Organization (2004). *The International Statistical Classification of Diseases and Health Related Problems ICD-10: Tenth Revision.* World Health Organization.

Wrase, J., Grüsser, S. M., & Heinz, A. (2006). Reizinduziertes Alkoholverlangen. *Der Nervenarzt, 77*(9), 1051-1063. https://doi.org/10.1007/s00115-006-2067-1

Wrase, J., Grüsser, S. M., Klein, S., Diener, C., Hermann, D., Flor, H., Mann, K., Braus, D. F., & Heinz, A. (2002). Development of alcohol-associated cues and cue-induced brain activation in alcoholics. *Eur Psychiatry, 17*(5), 287-291. https://doi.org/10.1016/s0924-9338(02)00676-4

Yalachkov, Y., Kaiser, J., & Naumer, M. J. (2012). Functional neuroimaging studies in addiction: multisensory drug stimuli and neural cue reactivity. *Neurosci Biobehav Rev, 36*(2), 825-835. https://doi.org/10.1016/j.neubiorev.2011.12.004

Zimmermann, U., Spring, K., Wittchen, H. U., & Holsboer, F. (2004). Effects of ethanol administration and induction of anxiety-related affective states on the acoustic startle reflex in sons of alcohol-dependent fathers. *Alcohol Clin Exp Res, 28*(3), 424-432. https://doi.org/10.1097/01.alc.0000117835.49673.cf

6 CURRICULUM VITAE

PERSONALIEN

Name:	Iris Reinhard
Wohnort:	Wintergasse 134, 69469 Weinheim
Geburtsdatum:	09.11.1962
Geburtsort:	Mannheim
Staatsangehörigkeit:	Deutsch
Kinder:	Kai Becker (22 Jahre)

SCHULISCHER WERDEGANG

1969 – 1972:	Grundschule Mannheim-Seckenheim
1972 – 1981:	Carl-Benz-Gymnasium Ladenburg
28.05.1981:	Abitur

UNIVERSITÄRER WERDEGANG

WS 1981/1982:	Beginn des Studiums (Mathematik, Diplom) mit Nebenfach Wirtschaftswissenschaften an der Universität Heidelberg
08.10.1984:	Vordiplom
28.05.1990:	Diplom
	Diplomarbeit: „Über das Konvergenzverhalten einiger populationsabhängiger Markoffscher Verzweigungsprozesse“
WS 1990/1991 – SS 1991:	Studium (Volkswirtschaftslehre, Diplom)

PUBLIKATIONSLISTE

https://www.zi-mannheim.de/forschung/personen/person-publikationen/person-publikationen/252/jahr/2021.html

7 DANKSAGUNG

Mein herzlichster Dank gilt vor allem meiner Doktormutter Frau Prof. (apl.) Dr. Sabine Vollstädt-Klein, die mir diese Doktorarbeit ermöglicht hat. Sie hat mich immer mit ihrem Rat und neuen Ideen unterstützt, und durch ihre hilfreichen Kritiken und ihr immenses Wissen im Bildgebungsbereich konnte die Arbeit stets verbessert werden. Auch ihr kollegialer Druck war nützlich, um diese Arbeit zu vollenden.

Meinen Koautorinnen und Koautoren danke ich für die effektive und freundschaftliche Zusammenarbeit, ohne die diese Studien nicht durchführbar gewesen wären. Hier möchte ich ganz besonders Frau PD Dr. Tagrid Leménager und Herrn Dr. Holger Hill danken, die mich immer wieder motiviert, gute fachliche Empfehlungen gegeben und auch kontroverse Diskussionen angeregt haben. Frau Sabine Hoffmann danke ich für die jahrelange Unterstützung im Datenmanagement- und Statistik-Bereich.

Ein großer Dank geht auch an meinen Chef, Herrn Prof. Dr. Stefan Wellek, der mir als Leiter der Abteilung Biostatistik die Möglichkeit eröffnet hat, mich für die Promotion in Richtung Suchtmedizin zu orientieren. Meinen Kolleginnen und Kollegen aus der Biostatistik möchte ich ganz herzlich danken für das jahrelange angenehme Arbeitsklima: Dr. Christine Jennen-Steinmetz, Peter Ziegler, Dr. Bertram Krumm und Mireille Lukas. Zudem sei Prof. Dr. Stefan Wellek und Frau Dr. Christine Jennen-Steinmetz für ihre Geduld und unzählige statistische Ratschläge gedankt.

Für den wertvollen Austausch bei vielen weiteren Projekten im Zentralinstitut für Seelische Gesundheit danke ich Frau Prof. (apl.) Dr. Christine Kühner, Herrn Prof. (apl.) Dr. Hans Joachim Salize, Frau Rosemarie Krämer, Frau PD Dr. Anne Koopmann, Herrn PD Dr. Patrick Bach und Herrn Prof. Dr. Markus Reichert.

Danke an alle Kolleginnen und Kollegen, Freundinnen und Freunde, die immer wieder leise nach dem Fortschritt der Arbeit gefragt und somit den Druck aufrechterhalten haben.

Ebenso danke ich Katja Engelke ganz herzlich für die intensive Durchsicht meiner Arbeit aus fachfremder Perspektive.

Ein besonderer Dank geht an meinen Lebenspartner Michel für die liebevolle Unterstützung in allen Lebenslagen, ebenso an unseren Sohn Kai, der durch die Wahl seines Studienfachs meine Motivation zur Fertigstellung dieser Arbeit weiter gefördert hat. Danken möchte ich auch meinen MitbewohnerInnen Julia und Leo, die meine Stimmungsschwankungen und zeitlichen Engpässe während dieser Zeit ins tägliche Miteinander integriert haben.

8 APPENDIX

Statement of own contribution

Work steps	**Publication 1**	**Publication 2**
Concept	Partly; for the survival analysis part completely	Partly; for the survival analysis part completely
Literature review	Partly; for the survival analysis part completely	Partly; for the survival analysis part completely
Data collection	n.a.	n.a.
Data analysis	Partly; for the survival analysis part completely	Partly; for the survival analysis part completely
Interpretation of results	Predominantly	Partly; for the survival analysis part completely
Preparation of manuscript	Predominantly	Partly; for the survival analysis part completely
Revision of manuscript	Predominantly	Partly; for the survival analysis part completely

Explanation of Equal Contribution in Publication 2:

The three equally contributing authors show expertise in separate areas in terms of content. The first author took care of the psychological aspects. Data processing of the raw data (artefact correction and estimation of the startle amplitudes) was performed by the second author. The author of this thesis (third equally contributing author) was responsible for all further statistical analyses of the pre-processed data. The preparation of the manuscript was shared according to the expertise of the three authors.