FROM DNA TO CBT
AN INVESTIGATION OF PSYCHOLOGICAL TREATMENT RESPONSE AND GENETICS

Evelyn Andersson

Stockholm 2018
Cover image: This pencil sketch of the DNA double helix was drawn by the British scientist Frances Crick in 1953. Crick and his American collaborator, James Watson, later visualized the structure of the DNA molecule through an iconic physical metal model that demonstrated how DNA replicates and how hereditary information is coded.


All previously published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet.
Printed by E-print AB 2018
© Evelyn Andersson, 2018
ISBN 978-91-7831-034-0
FROM DNA TO CBT
AN INVESTIGATION OF PSYCHOLOGICAL
TREATMENT RESPONSE AND GENETICS

THESIS FOR DOCTORAL DEGREE (Ph.D.)

Friday 1st of June, 2018. 2.00 PM

By

Evelyn Andersson

Principal Supervisor:
Associate Professor Christian Rück
Karolinska Institutet
Department of Clinical Neuroscience
Centre for Psychiatry Research

Co-supervisors:
Professor Nils Lindefors
Karolinska Institutet
Department of Clinical Neuroscience
Centre for Psychiatry Research
Professor Martin Schalling
Karolinska Institutet
Department of Molecular Medicine
and Surgery
Associate Professor Catharina Lavebratt
Karolinska Institutet
Department of Molecular Medicine
and Surgery
Associate Professor Erik Hedman-Lagerlöf
Karolinska Institutet
Department of Clinical Neuroscience
Division of Psychology

Opponent:
Professor Erik Jönsson
University of Oslo
Institute of Clinical Medicine

Examination Board:
Associate Professor Lars Westberg
University of Gothenburg
Department of Neurosciences and Physiology
Associate Professor Lisa Thorell
Karolinska Institutet
Department of Clinical Neuroscience
Associate Professor Mia Ramklint
Uppsala University
Department of Neuroscience
For my grandpa Evert.
ABSTRACT

Background. Social anxiety disorder, panic disorder, and major depressive disorder are common, disabling, and heritable disorders. The first line of treatment for all anxiety disorders and mild to moderate depression is cognitive behavior therapy (CBT). However, 40–60% of patients fail to respond adequately to treatment and the ability to a priori predict who will respond to treatment is essential for being able to choose alternative treatments to those who will fail the treatment. As clinical and demographic variables so far have shown limited predictive value of in guiding therapeutic decisions, it has been suggested to incorporate genetic variation in the set of predictors to increase precision.

Aims. The general aim of this thesis was to investigate the genetic underpinnings of psychological treatment outcomes for social anxiety disorder, panic disorder, and major depressive disorder. Specifically, we aimed to identify the clinical and genetic predictors of patients with social anxiety, and to evaluate participant responses to CBT (Study I); next we increased the sample size from Study I and collaborated with another site and investigated three polymorphisms and their predictive value in CBT response for social anxiety (Study II). Subsequently, we aimed to calculate polygenic risk scores by performing a genome-wide association analysis to study genetic variation and CBT responses to major depression (Study III); finally, we further extended the sample size by performing a meta-analysis, and by calculating polygenic risk scores to study potential genetic overlap with other cognitive and psychiatric traits, and the link between genetic variations and CBT responses in patients with anxiety disorders.

Methods. In Study I and II, we recruited participants from randomized controlled trials to investigate clinical variables (Study I) and candidate gene polymorphisms (Study I and II); as potential predictors of CBT response for social anxiety disorder. Study III and IV we recruited larger samples and performed genome-wide association analyses to estimate genetic variance in CBT response and to calculate polygenic risk for CBT outcome for major depression and anxiety disorders.

Results. In Study I, several clinical, but neither of the genetic variables predicted CBT outcome for social anxiety. In Study II, neither of the genetic polymorphisms predicted response to CBT for SAD. In Study III, an association of higher load of autism spectrum polygenic risk scores and less decrease in depressive symptoms after CBT was shown. In Study IV, no genome-wide significant loci or genetic overlap with psychiatric or cognitive traits were present.
Conclusions. The results in this thesis present the first findings of an association of aggregated genetic risk scores and CBT outcome in depression. Our results indicate that the use of polygenic risk scores may be a fruitful approach in genetic studies of CBT outcomes. In addition, the results provide support for the continued efforts of collecting larger and more homogenous samples in studies of complex traits such as psychological treatment response.
LIST OF SCIENTIFIC PAPERS


CONTENTS

1 Introduction.................................................................................................................. 1
  1.1 Anxiety disorders and major depression ............................................................... 2
      1.1.1 Social anxiety disorder (SAD) ........................................................................ 2
      1.1.2 Panic disorder (PD) .................................................................................... 2
      1.1.3 Major depressive disorder (MDD) .............................................................. 3
  1.2 Cognitive behavior therapy (CBT) ........................................................................ 3
      1.2.1 Avoidance and exposure ........................................................................... 4
      1.2.2 Internet-based CBT (ICBT) ....................................................................... 4
  1.3 Non-genetic predictors of CBT .............................................................................. 5
  1.4 DNA and genetic association methods ................................................................. 6
      1.4.1 DNA overview and collection ..................................................................... 6
      1.4.2 Candidate gene approach .......................................................................... 8
      1.4.3 GWA studies ............................................................................................. 8
      1.4.4 Polygenic risk scores ................................................................................. 10
  1.5 Anxiety disorders/depression and genetics ......................................................... 11
      1.5.1 Polymorphisms of interest in anxiety disorders and depression .............. 11
      1.5.2 Recent genetic findings relating to anxiety disorders and depression .... 12
  1.6 Therapygenetics ..................................................................................................... 13
  1.7 Aims ....................................................................................................................... 15

2 Methods ....................................................................................................................... 17
  2.1 Participants ........................................................................................................... 17
      2.1.1 The Karolinska Institutet sample (Study I–IV) ........................................... 17
      2.1.2 Study sample from Uppsala University (Study II) ...................................... 17
      2.1.3 Study sample from UK and Germany (Study IV) ..................................... 18
  2.2 The intervention .................................................................................................... 18
      2.2.1 ICBT (Study I–IV) .................................................................................... 18
      2.2.2 Group-delivered CBT (GCBT) (Study I and IV) .................................... 21
      2.2.3 Standard CBT (Study IV) ......................................................................... 22
  2.3 Primary outcome measures ................................................................................... 22
      2.3.1 Liebowitz Social Anxiety Scale (LSAS) ..................................................... 22
      2.3.2 Self-rated Montgomery Åsberg Depression Rating Scale (MADRS) ...... 22
      2.3.3 Panic Disorder Severity Scale (PDSS) ....................................................... 22
      2.3.4 Clinical Global Impression-Severity (CGI-S) ......................................... 23
      2.3.5 Panic Agoraphobia Scale (PAS) ................................................................. 23
  2.4 Definition of treatment responses ........................................................................... 23
      2.4.1 Study I and II ............................................................................................ 23
      2.4.2 Study III .................................................................................................... 23
      2.4.3 Study IV .................................................................................................... 24
  2.5 Collecting genotypes .............................................................................................. 24
2.5.1 DNA collection .............................................................................. 24
2.6 Statistical analysis ............................................................................. 24
  2.6.1 Study I and II ............................................................................... 24
  2.6.2 Study III ..................................................................................... 24
  2.6.3 Study IV ..................................................................................... 25
2.7 Genotyping ....................................................................................... 25
  2.7.1 Study I ...................................................................................... 25
  2.7.2 Study II ..................................................................................... 25
  2.7.3 Study III .................................................................................... 26
  2.7.4 Study IV ..................................................................................... 26
  2.7.5 Quality control ........................................................................... 26
  2.7.6 Principal components ancestry (PCA) ......................................... 27
  2.7.7 Polygenic risk scores (PRS) ......................................................... 27
2.8 Ethical considerations ......................................................................... 28
3 Empirical Studies ............................................................................... 29
  3.1 Study I .......................................................................................... 29
  3.2 Study II ........................................................................................ 29
  3.3 Study III ......................................................................................... 31
  3.4 Study IV ......................................................................................... 32
4 General Discussion ............................................................................ 36
  4.1 Can clinical variables and candidate gene polymorphisms predict CBT
treatment response in SAD? ................................................................. 36
  4.2 Can we detect genetic predictors of CBT response in SAD if we double the
sample size through a multi-center approach? ...................................... 37
    4.2.1 Methodological considerations in Study I and II ......................... 38
  4.3 Can aggregated genetic scores based on genome-wide association studies
on psychiatry-related traits predict treatment response in MDD? .......... 38
    4.3.1 Methodological considerations for Study III ............................... 39
  4.4 Can we estimate genetic variance in treatment response from a genome-
wide study with twice the sample size as that of Study III? Can we
identify genetic overlap of CBT response and phenotypes for psychiatric
and cognitive traits? ............................................................................. 40
    4.4.1 Methodological considerations for Study IV ............................... 41
  4.5 Future directions ............................................................................. 41
  4.6 Ethical concerns for future genetics studies .................................... 43
5 Conclusion .......................................................................................... 44
6 Acknowledgements ............................................................................. 45
7 References .......................................................................................... 49
LIST OF ABBREVIATIONS

ACC  Anterior cingulate cortex
ADHD Attention deficit/hyperactivity disorder
AG  Agoraphobia
ANOVA Analysis of variance
ASD Autism spectrum disorder
BDNF Brain-derived neurotrophic factor
BIP Bipolar disorder
CBT Cognitive behavior therapy
CGI-S Clinical global impression-severity
CIDI Composite International Diagnostic Interview
CNV Copy number variant
COMT Catechol-O-methyltransferase
DIPS Diagnostisches Interview bei Psychischen Störungen
EDU Educational attainment
fMRI Functional magnetic resonance imaging
GCBT Group-delivered cognitive behavioral therapy
GENDEP Genome-based therapeutic drugs for depression
GRS Genetic risk score
GWA Genome-wide association
HWE Hardy-Weinberg Equilibrium
ICBT Internet-based cognitive behavioral therapy
IQ Intelligence quotient
LD Linkage disequilibrium
LSAS-SR Self-rated Liebowitz social anxiety scale
MADRS-S Self-rated Montgomery Åsberg depression rating scale
MCC Midcingulate cortex
MDD Major depressive disorder
MINI Mini-International Neuropsychiatric Interview
MLMA Mixed linear model association
PAS Panic agoraphobia scale
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC</td>
<td>Principal component</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal component ancestry</td>
</tr>
<tr>
<td>PD</td>
<td>Panic disorder</td>
</tr>
<tr>
<td>PDSS-SR</td>
<td>Self-rated panic disorder severity scale</td>
</tr>
<tr>
<td>PFC</td>
<td>Prefrontal cortex</td>
</tr>
<tr>
<td>PGC</td>
<td>Psychiatric Genomics Consortium</td>
</tr>
<tr>
<td>PRS</td>
<td>Polygenic risk score</td>
</tr>
<tr>
<td>PTSD</td>
<td>Post-traumatic stress disorder</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomized controlled trial</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver operator characteristics</td>
</tr>
<tr>
<td>SAD</td>
<td>Social anxiety disorder</td>
</tr>
<tr>
<td>SCID</td>
<td>Structured Clinical Interview for DSM Disorders</td>
</tr>
<tr>
<td>SCZ</td>
<td>Schizophrenia</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SP</td>
<td>Specific phobia</td>
</tr>
<tr>
<td>SRI</td>
<td>Serotonin reuptake inhibitor</td>
</tr>
<tr>
<td>VMPFC</td>
<td>Ventromedial prefrontal cortex</td>
</tr>
<tr>
<td>VNTR</td>
<td>Variable-number tandem repeat</td>
</tr>
<tr>
<td>5-HT</td>
<td>Serotonin</td>
</tr>
<tr>
<td>5-HTT</td>
<td>Serotonin transporter protein</td>
</tr>
<tr>
<td>5-HTTLPR</td>
<td>Serotonin-transporter-linked polymorphic region</td>
</tr>
</tbody>
</table>
1 INTRODUCTION

“For every complex problem there is an answer that is clear, simple, and wrong.”
—H. L. Mencken

Attending a clinical psychology program in 2000 provided insights into many aspects of the complex human mind, the continuous pursuit of ways to reduce the burden of mental illnesses, and the long-standing questions of both clinicians and patients—why, how, and why me. Personally, it came as a surprise to realize just how deep the nature versus nurture controversy went, and how this dispute of the 20th century still vividly flourished. This controversy is an argument over the extent to which two different sources influence the development of behavior: biological disposition (genes) or environmental experience (upbringing). The answer to the questions of why, how, and why me are not only difficult to answer, but rather the mere question itself segregated the whole field of psychology with different schools advocating for their own perspective. Soon, these experiences left me impatient, with a desire to attempt to understand more of what, at first glance, looked like a cruel lottery deciding who was burdened with mental illness, the patient’s individual prognostic outlines, and which psychological treatments would be most likely to work for each patient. Given the vast evidence supporting a multimodal understanding of human behavior, it is hard to imagine someone today having a categorical view on psychiatric etiology and prognosis. However, when the work on this thesis started in 2012, the psychiatric genetics field had been focusing on the discovery of a few critical genes involved in mental illness for the previous decade. This later turned out to be harder than expected, and the collaborative effort of collecting enormous samples to identify common variants across the genome was in progress; finally; the breakthroughs began. With larger samples sizes came more conclusive findings. Not long after, the question of the utility of the findings was emphasized. The beginning of the search for answers to my itching questions had begun—did treatment response relate to a person’s individual biological makeup, and could genes play a part in the response of a patient? These simple questions—what works for whom and why—are accompanied by not so simple answers.

Stockholm, April 2018
1.1 ANXIETY DISORDERS AND MAJOR DEPRESSION

Anxiety disorders and depression are disorders with among the highest global burden of disease, as measured by years lived with the disability.\(^1,2\) Panic disorder (PD), social anxiety disorder (SAD), and major depressive disorder (MDD) are highly prevalent; over half the population is affected by at least one of them at some point during their lives.\(^3\) The reasons for the large burden of disease associated with these disorders include their early onset, chronicity, recurrent nature, the lack of access to treatment, and limited efficacy of treatments.\(^1,3,4\)

1.1.1 Social anxiety disorder (SAD)

SAD is characterized by extreme discomfort and avoidance of social situations in which the individual is subject to scrutiny by others.\(^5\) SAD has a lifetime prevalence of 10\%, a chronic course, and an early age of onset.\(^1\) Furthermore, it is associated with functional impairment in several areas. Half of the population with SAD has been shown to meet the criteria for another psychiatric disorder sometime during their lives; depression- and other anxiety-related disorders are the most common.\(^6\)

1.1.2 Panic disorder (PD)

PD is associated with intense attacks of anxiety, often followed by avoidance behavior.\(^7\) It has a lifetime prevalence of around 5\%, and the lifetime prevalence for isolated panic attacks has been estimated at 22.7\%.\(^8\) Individuals with PD experience recurrent, unexpected panic attacks accompanied by a range of physiological and cognitive signs such as an elevated heart rate, sweating, difficulty breathing, depersonalization, and a debilitating fear that they are dying or going crazy. This intense experience of terror leads to a fear of additional attacks and results in extensive avoidance behavior. PD is often followed by agoraphobia—the anxiety of being trapped or captured somewhere from which escape will be difficult or where help will be unavailable when a panic attack occurs. PD with or without agoraphobia is highly comorbid with other mental disorders, particularly other anxiety-related disorders.\(^3\)
1.1.3 Major depressive disorder (MDD)
Depression is characterized by persistent low mood, fatigue, hopelessness, concentration issues, morbid thoughts of death, feelings of guilt, and a range of other symptoms. Moreover, MDD is often associated with both personal and economic problems. The lifetime prevalence of MDD has been estimated at between 19.8% in low-income countries and 28.1% in high-income countries.

1.2 COGNITIVE BEHAVIOR THERAPY (CBT)
CBT is an umbrella term for short-term psychological therapy that includes both cognitive and behavioral aspects. CBT has strong empirical support for treating mild to moderate depression and anxiety disorders, and is recommended as the first line of treatment by international and national guidelines. CBT for PD, SAD, and MDD (mild to moderate) have demonstrated moderate to strong effect sizes in hundreds of randomized controlled trials (RCTs) and meta-analyses. Although there is a wide variety of different CBT treatment manuals and interventions, several core components are usually part of the CBT repertoire. First, a focus of CBT is identifying and changing maladaptive thought patterns such as negative recurring thoughts through a process referred to as cognitive restructuring. Second, learning about and addressing emotions and aversive feelings is a primary component of CBT manuals. Distorted thoughts and cognitions leads to problematic emotions and behavior (avoidance). Finally, behaviors are believed to exacerbate problematic patterns, and CBT aims at disrupting the cognitive cycle of negative behavior by confronting negative thoughts and cognitions by seeking out the validity in them. This spiral of reinforcing behavior is shown in Figure 1, and the CBT therapist addresses all of them during treatment.

![Figure 1. The cognitive cycle of behavior.](image-url)
1.2.1 Avoidance and exposure

In CBT for any anxiety disorder, graded exposure to fear-inducing situations is an important intervention. When people are fearful of something, they tend to avoid the feared object, disregarding their intellectual knowledge that the feared object is harmless.\(^{15}\)

Avoidance reduces feelings of fear in the short term, but unintentionally sustains and even increases the fear in the long-term. Exposure is most effective when it is predictable, frequent, and prolonged.\(^{16}\) This ICBT program for PD involves two types of exposure: *interoceptive exposure* and *in vivo exposure*.

*Interoceptive exposure* is gradual exposure to internal physical sensations of anxiety. For instance, a patient suffering from PD typically fears the feeling of an elevated heart rate. The increase in heart rate resembles the sensation of the onset of a panic attack, and therefore the patient tries to avoid any behavior that can trigger an increased heart rate, such as drinking coffee or physical exercise. In an interoceptive exposure exercise, the individual will gradually induce sensations associated with these fears—for example, consume caffeine—and rate their level of anxiety until the fear subsides.

*In vivo exposure* means directly facing an external feared stimulus. The patient faces a feared object, situation, or activity in a natural setting. For example, someone with a fear of riding the subway might be instructed to practice that activity in a predictable and gradual manner.

1.2.2 Internet-based CBT (ICBT)

ICBT is an umbrella term for digitalized therapy, and ranges from automated e-learning programs to computer-based aid with traditional CBT. ICBT is a format that provides access to treatment regardless of geographic location or time, through the internet.\(^{17}\) Over the past two decades, a substantial body of evidence has suggested that ICBT is an efficacious treatment for MDD and anxiety disorders.\(^{18}\) A meta-analysis of over twenty studies has demonstrated has comparable effects and outcomes of ICBT to those of traditional face-to-face CBT.\(^{18}\)
1.3 NON-GENETIC PREDICTORS OF CBT

Although CBT is the first treatment of choice for both mild to moderate depression and anxiety disorders,\textsuperscript{9, 12} about 30–60\% of patients undergoing CBT will not respond sufficiently to this treatment.\textsuperscript{19} Finding robust predictors of CBT outcomes would enable the personalization of treatment choice.\textsuperscript{20, 21} A number of factors have been reported as potential predictors of a poor treatment outcome in CBT, though many of them are somewhat inconsistent.\textsuperscript{22}

**Socio-demographic predictors:** Socio-demographic factors related to a beneficial CBT outcome are for instance, numbers of academic years, strong social support\textsuperscript{23} and working full-time.\textsuperscript{24} However, the direction of the CBT outcome is conflicted between studies; in one study, the level of education predicted a faster recovery,\textsuperscript{25} whereas another study suggested that this predicted a decrease in therapy response.\textsuperscript{26, 27} One recent meta-analysis, which included fifty-two studies into the predictors of CBT for PD, rejected the previous findings of socio-demographic relevance on the outcome.\textsuperscript{22}

**Clinical characteristics as predictors:** One of the most studied clinical predictors of CBT is the baseline severity of the illness. For instance, in a recent large meta-analysis of 1,201 clinical cases of patients with anxiety or depression disorder who had completed an ICBT program, initial symptom severity had a clear association with treatment outcomes.\textsuperscript{28} This was contrary to previous findings, however, where baseline severity was linked to a poorer treatment response.\textsuperscript{29} In another study investigating the initial choice between medication or CBT, the severity of depression did not predict the outcome of either treatment strategy.\textsuperscript{30} Furthermore, a recent meta-analysis of fifty-two reports on CBT response showed that a comorbid diagnosis decreased treatment response.\textsuperscript{30}

**Treatment-related predictors:** Prediction studies of ICBT have given extra attention to treatment-related factors such as treatment credibility; this is because the nature of the ICBT format resembles self-help programs to some extent, and requires patient adherence and self-motivation to a larger extent.\textsuperscript{24} In line with this knowledge, treatment credibility has been reported as a significant factor associated
with therapy response. In addition, adherence also seems to be a predictor relevant to CBT outcome, both in traditional CBT and in ICBT.

**Neurobiological predictors:** Functional magnetic resonance imaging (fMRI) can be used to study changes in the brain during behavioral or cognitive tasks. In a study of treatment responses to PD, hyperactivation of the “fear network,” including the brainstem, anterior- and mid-cingulate cortex (ACC and MCC), insula, and part of the prefrontal cortex (PFC), was demonstrated. Also, in SAD responders of the long-term outcome of CBT, an abnormal response pattern in the ACC and the amygdala was demonstrated. In addition, in a study of patients with SAD, hypoactivation of the ventromedial PFC (vmPFC) and hyperresponsivity in the amygdala predicted CBT treatment outcome. Generally, the neurobiological aspects of CBT response were correlated with areas involved in regulating affective states and general emotional processing, such as the amygdala, the insular cortex, the ACC, and parts of the PFC.

In summary, the results of non-genetic CBT response tests are mixed, often based on clinical trials with disparate outcome measures, and have so far not been able to reach an acceptable predictive power to guide clinical decisions. The majority of predictor studies focused on non-genetic factors; given the expected heritability of any complex behavior (especially in terms of treatment response), genetic markers ought to be investigated in the pursuit of better treatment outcome predictions.

### 1.4 DNA AND GENETIC ASSOCIATION METHODS

#### 1.4.1 DNA overview and collection

Genetic association studies aim to identify risk alleles related to factors such as a disease or a behavioral trait—such as treatment response. The basis of this association depends on the assumption that genes are directly involved in the development and function of brain regions involved in behavior, their neural pathways, and the effects of functional gene variants on these processes. Single nucleotide polymorphisms (SNPs) are the most common targets of association studies, but there are a range of other less frequent markers, such as insertion/deletions, variable-number tandem repeats (VNTRs) and copy-number variants (CNVs). DNA is usually sampled by extracting cells from a blood sample, a cheek
swab, or a spit test. Next, the DNA is extracted from the cells and a solution with an allele marker and fluorescent dye are added. If the allele of interest is present in the sample, the marker will bind to that gene and activate the dye, thereby allowing researchers to observe the genotype (genetic trait) of that person.

**Figure 2. DNA overview.**
1.4.2 Candidate gene approach

Before genome-wide association (GWA) arrays (i.e., the genotyping chip) became widely available, candidate gene approaches dominated the field.\(^{38}\) The rationale of the candidate gene approach is that a major component of the quantitative genetic variations of the trait under investigation is caused by a functional mutation in the gene of interest (the candidate gene).\(^{39}\) This link is established by association, meaning that a certain genotype is associated with a certain phenotype, and the loci of interest are assumed to be causally related to the observed trait. The process of selecting relevant candidate genes is usually guided by previous studies and databases with functional genomic data.\(^{40}\) Candidate gene studies are less expensive than other genotyping methods, and they are focused on the selection of genes with previously known direct or indirect biological functions related to the regulatory process of the disease or trait. Eventually, the association between the genotype and the trait is usually verified through a case-control study by observing the frequency of the variant in random cases and comparing them to the frequency of the gene variant in random controls.\(^{37}\) Thus, the design rests heavily on an \textit{a priori} hypothesis of a critical gene in causal biological pathways that is directly associated with the trait of interest. However, a minority of diseases are monogenic—caused by a functional mutation in a single gene, where a malfunctioning gene affects the production of a crucial protein.\(^{36}\) An example of a monogenic disorder is Huntington’s disease. In addition to the fact that psychiatric disorders in general are polygenic, even in the presence of highly plausible biological hypotheses the risk of selecting a candidate gene with infinite small effect size in too small samples (underpowered samples) are present, which in turn increases the risk of both false-positives and false-negatives. In fact, a minority of the initial candidate gene findings were later replicated.\(^{41}\) However, some have argued that these many discouraging results in candidate gene studies were not a complete failure, as they forced the field into more appropriately powered studies and a focus on small-effect common variants.\(^{42}\)

1.4.3 GWA studies

In GWA studies, a large number (up to 1.1 million) of markers for common genetic variants (SNPs) are examined to analyze genetic variation across the entire genome (Figure 3). As in other association studies with case-control designs, GWA studies rely on the fact that a variant leading to a phenotype is found at a higher frequency in cases (individuals with the phenotype) than in controls (individuals without the phenotype).\(^{43}\) Statistical analysis is performed to indicate the probability and effect size of the specific genotype to be associated with the trait. In contrast to the candidate gene approach, the GWA method is
hypothesis free, resulting in an unbiased framework for analyzing the predictive power of each of one million individual SNPs. Because so many statistical tests are performed in a GWAS (> 1,000,000), the risk of false-positive outcomes is great. To counteract the risk of multiple testing errors, the significance threshold is set at $p < 5 \times 10^{-8}$, rather than the widely used value of $p < 0.05$ that is standard in single tests. In addition, given the strict p-value thresholds, in order to have enough power to detect signals from genetic variants with small effect sizes, a large sample size is crucial.\textsuperscript{44} However, with a sufficient sample size, GWA studies can reveal important etiological and biological insights. Examples include the identification of hundreds of loci for diseases such as inflammatory bowel disease, type 2 diabetes, multiple sclerosis, and rheumatoid arthritis. The Swedish Schizophrenia Study Identified over 200 significant loci and found that common SNPs collectively accounted for > 50% of heritability for schizophrenia (SCZ).\textsuperscript{45} At present, GWA studies have identified specific loci for SCZ, bipolar disorder (BIP),\textsuperscript{46} attention deficit/hyperactivity disorder (ADHD),\textsuperscript{46} ASD\textsuperscript{47} and most recently, MDD.\textsuperscript{48} In addition, some genes were found to be associated with different psychiatric phenotypes, suggesting that there may be several general psychiatric vulnerability genes.\textsuperscript{49} Furthermore, a recent study found a strong genetic correlation between psychiatric disorders (ASD, ADHD, Tics disorder, obsessive compulsive disorder, anxiety, MDD and schizophrenia) and continuous related population traits, indicating that psychiatric disorders are at an extreme end of a continuum in the normal population.\textsuperscript{50} The first GWA study of treatment outcome for anxiety disorders (n = 980) did not detect any significant common variants, and experience from other fields would suggest that larger sample sizes are needed.\textsuperscript{51} With the focus on smaller-effect variants, the psychiatric genetics field finally gained some insight into the genetic architecture of mental illnesses, as well as an essential understanding of behavioral and psychiatric traits as highly polygenic.\textsuperscript{42}
Figure 3. A genome-wide association (GWA) study. The genetic variants of interest were more frequent in cases than in controls. A GWA chip or SNP array is a DNA microarray tool for high-throughput analysis of SNPs. Modern SNP arrays (used in Study III and IV) can analyze up to 1.1 million SNPs in a single chip. The SNP associations are presented in a Manhattan plot, a scatter plot with visual resemblance to the Manhattan skyline, where each dot represents the \( p \)-value for each SNP association on the Y-axis, and the genetic positions (loci) on the X-axis. The spikes or “skyscrapers” that reach above the dotted horizontal line are loci with GWA significance.

1.4.4 Polygenic risk scores

GWA studies aim to find genetic loci associated with a phenotype (a disease or treatment outcome). Given that the contribution to the explained variance of each individual SNP is
small, one way forward is to calculate a risk score based on multiple loci. Genetic or polygenic risk scores (GRS or PRS) are based on the idea that the number of risk alleles carried by an individual measures genetic liability.\textsuperscript{52} GRS are calculated by adding the risk alleles at a given GWA study \textit{p}-value threshold. The more risk alleles an individual has, the higher the GRS. The potential utility of GRSs was demonstrated in a study where a high GRS for coronary heart disease predicted other disease-related risks, and subsequently was used for individualized therapeutic interventions.\textsuperscript{53} The “risk” label of GRS/PRS do not necessarily indicate a risk per se; the scores are distributed normally for both adverse and beneficial outcomes, such as having a genetic “risk” for qualifying for high intelligence.

\section*{1.5 ANXIETY DISORDERS, DEPRESSION AND GENETICS}

\subsection*{1.5.1 Polymorphisms of interest in anxiety disorders and depression}

At the start of this thesis, there were promising findings of candidate gene variants associated with SAD, PD, and depression.\textsuperscript{54, 55} One of the most frequently studied gene variants was an insertion/deletion mutation in the serotonin transporter gene promoter (5-HTTLPR).\textsuperscript{55} Serotonin (5-HT) is a central nervous system neurotransmitter, involved in regulating psychological (e.g., fear, learning, etc.)\textsuperscript{56} and physical functions (e.g., appetite or sleep).\textsuperscript{57} The serotonin transporter protein (5-HTT), encoded by the \textit{SLC6A4} gene, is involved in transporting serotonin from the synaptic space and in facilitating its reuptake to the presynaptic neuron. The \textit{SLC6A4} gene has a variant that modulates the transcription of the gene: the 5-HTT linked polymorphic region (5-HTTLPR); its repetitive sequence varies in length (either a long [L] or a short [S] allele). Many psychotropic medications target 5-HTT, especially serotonin reuptake inhibitors (SRIs).\textsuperscript{58} More than twenty-five studies have shown that allelic variations in 5-HTTLPR are associated with characteristics (neuroticism) relevant to SAD.\textsuperscript{59} Around year 2008, when Study I and II of this thesis were planned, studies linked the 5-HTTLPR to social anxiety,\textsuperscript{60} and wide range of studies showed an association for the 5-HTTLPR and PD,\textsuperscript{61} as well as with depression.\textsuperscript{62} However, few of these findings were replicated,\textsuperscript{63} and a meta-analysis of eleven different cohorts concluded that the findings of association of 5-HTTLPR and anxiety were not reliable or replicable.\textsuperscript{64}

Another polymorphism hypothesized to be of relevance is in the gene encoding the Catechol-O-methyltransferase (COMT) enzyme. COMT is involved in a broad range of biological functions, such as regulation of catecholamines, and the \textit{COMT} gene encodes a protein expressed in the brain—membrane-bound COMT—that is responsible for the modulation of dopamine.\textsuperscript{65} A functional polymorphism in COMT, V158M, has been
associated with many psychiatric conditions.\textsuperscript{66, 67} Several studies have also suggested that the V allele of the COMT V158M mutation is associated with impaired emotion perception (difficulties reading and/or interpreting emotional signals),\textsuperscript{68, 69} a process highly relevant in SAD and PD, and perhaps also in treating them.\textsuperscript{70} For instance, a person with SAD typically has difficulties estimating the intentions of others; in a CBT for SAD, behavioral experiments (see section 2.2.1.1)—especially planned activities to test out negative thoughts in action in everyday social situations—could be affected by the patient’s ability to perceive emotional information. The first study ever investigating genetics and the response to CBT found that the COMT V158M polymorphism was associated with treatment outcomes during the exposure phase of CBT in patients with PD.\textsuperscript{70}

Furthermore, the TPH2, an enzyme involved in modulating serotonin is encoded by the \textit{TPH2} gene. TPH2 regulates the rate-limiting synthesis of neuronal serotonin. TPH2 has been linked to emotional regulation by affecting brain regions involved in fear (acquisition and extinction) and anxiety,\textsuperscript{71-73} such as the amygdala; and has shown a direct association with social anxiety\textsuperscript{74} and depression.\textsuperscript{54} Given the central process of fear conditioning that is associated with anxiety disorders, and the fear extinction processes involved in treating both depression and anxiety disorders, the genetic variations involved in these processes could be of importance as possible predictors.

Another variant assumed to play a role in fear and emotional processing is the brain-derived neurotrophic factor (BDNF).\textsuperscript{75} BDNF is involved in neuronal survival and synaptic plasticity. BDNF is highly expressed in the brain, and a common SNP (V66M) in the human \textit{BDNF} gene is associated with the irregular intracellular trafficking of pro-BDNF.\textsuperscript{76} In a meta-analysis of BDNF and anxiety disorders, reduced BDNF levels showed an association with anxiety disorders.\textsuperscript{77}

### 1.5.2 Recent genetic findings relating to anxiety disorders and depression

The risk of developing social anxiety is higher in individuals with a close relative with SAD, suggesting a heritable basis to SAD.\textsuperscript{78} In addition, a recent GWA meta-analysis of social anxiety in different ancestral samples aimed to determine genetic loci for SAD found one locus each in the European and African ancestral samples on chromosome 6 and chromosome 1, respectively. These are the first genome-wide loci found in SAD.\textsuperscript{79}

Although the pathogenesis and neurobiological mechanisms of PD are not yet fully understood, some markers have been identified and are considered to have a strong link to the disorder.\textsuperscript{80} For instance, alterations in the amygdala pathway and irregularities in the
hypothalamic-pituitary-adrenal axis have been studied and linked to PD in several studies. Twin studies suggest heritability estimates of 30–40% for PD, and the genetic overlap across different anxiety disorders, such as agoraphobia and panic disorder, is considerable. GWA studies investigating PD have not been able to establish any genetic loci associated with PD. However, functional data recently demonstrated a variation in the *TMEM132D* gene in patients with PD. A large meta-analysis of candidate gene studies of PD examined twenty-three genetic variants in twelve putative susceptibility genes related to anxiogenesis. After correction for multiple testing, only two genes (*COMT* and *TMEM132D*) and three polymorphisms related to those genes (*TMEM132D* rs7370927 [T allele] and rs11060369 [C allele], and *COMT* rs4680 [G allele]), were significant in a European sample. This implies that the variants may be ancestry-specific, as this significance was not observed in non-European samples.

There is evidence of a genetic component to MDD, as heritability of the disorder is estimated at 37%. Both linkage studies and candidate gene approaches have yielded hundreds of conflicting results. The genome-wide approach also had to overcome many disappointing null findings, exemplified by a meta-analysis of GWA MDD (9,240 MDD cases and 9,519 controls) where no significant genome-wide findings were observed. Finally, the awaited breakthrough was made in a study involving a large enough sample size (138,000 cases) to detect significance, and forty-four new loci for MDD were reported. The genetic findings were associated with brain regions including the PFC and the ACC, areas associated with MDD. The fact that such a large sample size was needed to produce meaningful results speaks to the complexity and heterogeneity of MDD, given that smaller studies into other psychiatric disorders have been far more successful.

1.6 THERAPYGENETICS

The majority of psychiatric genetics research has focused on the genetic architecture of psychiatric disorders, but another research question is the genetic underpinnings of treatment outcomes. If we could identify genetic markers before the start of a certain treatment, this could lead to personalized alternative treatment choices, avoiding treatment failure, and a better allocation of limited health care resources. Pharmacogenetics aims to more clearly understand how patients respond differently to drugs based on genetic variation. A number of studies have attempted to investigate the predictive power of gene variants to which medications will be most effective for a given patient. Predicting the pharmacogenetic response for psychototropic medications (e.g., antidepressants) has until
now yielded few results. However, recently, the independently funded Genome-Based Therapeutic Drugs for Depression (GENDEP) used statistical learning models to identify common polymorphisms in 280 participants receiving antidepressants (escitalopram or nortriptyline) during a 12-week treatment period. In the statistical model twenty genetic variants predicted remission, and explained approximately 36% of variance in treatment response. In other disciplines of medicine, examples of discoveries in pharmacogenetics have led to mandatory genetic testing in advance of selecting drugs, dose, and duration.

The first pilot study to investigate genetic variation and response to psychological therapy found an association between COMT V158M and symptom relief in CBT for PD. Soon afterwards, the first study of the association between 5-HTTLPR and response to treatment for anxiety in children reported a more favorable CBT outcome if the individual had the homozygous short-short (S/S) genotype, rather than the heterozygous short-long (S/L) genotype. Previously, Caspi et al. (2003) had reported a gene-environment interaction in the same polymorphism, with others later suggesting that carriers of the homozygous short (S/S) had a susceptibility variant, rather than a vulnerability variant. This meant that having the S/S variant increased the likelihood of a better response to environmental interventions, such as treatment interventions. However, the result called for larger samples and the authors coined the term “therapygenetics” for the new field of psychological therapy response and genetics. A number of studies tried to replicate the initial findings of the 5-HTTLPR polymorphism, but showed mixed results. At the time, many polymorphisms were suggested targets of investigation; for instance, the TPH2 G703T mutation and the BDNF V66M mutation, but none could be systematically replicated.

Development in the genetics field spurred therapygenetics to sideline the candidate gene methodology, instead moving the focus to genome-wide approaches. GWA studies have become more widely available through collaborative efforts and a dramatic price drop in the price of the technology. However, only two GWA studies have been published on psychological treatment response to date. The first GWA study of response to psychological therapy was in children with anxiety disorders (n = 939), but no significant genetic associations with treatment response were observed, although four independent loci were potentially significant (p < 5 × 10^-6) at post treatment. A second treatment response GWA study (n = 102) detected three loci post-treatment, and four loci at six months follow-up that were also potentially significant (p < 5 × 10^-6). These results are in line with GWA findings in other areas, suggesting that complex behaviors, as therapy response is assumed to be, are highly polygenic and require larger sample sizes.
1.7 AIMS

The overall aim of this thesis was to relate genetic variations to psychological therapy outcome for PD, SAD, and MDD.

The specific aims of each study are presented below:

Study I: To identify clinical and genetic predictors for patients with SAD and participant response to CBT.

Study II: To investigate three candidate gene polymorphisms and their relationships with CBT response in SAD in a sample from two clinical sites.

Study III: To generate GRSs from six psychiatric and cognitive traits to test their predictive utility for ICBT response in MDD.

Study IV: To estimate the variance in treatment response to CBT for anxiety that could be explained by common genetic variants (SNP heritability) and to calculate polygenic scores to examine the genetic overlap of treatment response with psychiatric and cognitive traits.
2 METHODS

2.1 PARTICIPANTS

2.1.1 The Karolinska Institutet sample (Study I–IV)

The sample comprised adult patients with SAD, PD, and MDD who started group CBT (GCBT; Study I and IV) or ICBT at the Psychiatric Anxiety Clinic or the Internet Psychiatry Clinic in Stockholm; the latter is a psychiatric clinic specializing in delivering psychologist-guided ICBT. Patients in both groups (GCBT and ICBT) were asked to participate in the study, and had either been referred to the clinic by their general practitioner or via an online self-referral system. In both groups, after online screening, the patients came to the clinic for psychiatric assessments, including a structured diagnostic interview—either the Mini-International Neuropsychiatric Interview (MINI), or the Structured Clinical Interview for DSM Disorders (SCID)-I Research version (Study I)—which was used to establish diagnosis.

A psychiatrist or supervised resident performed the interview. For enrollment in the study, the patient had to meet the following requirements: fulfill the criteria in the DSM IV-TR for current SAD, PD, or MDD diagnosis; be able to read and write in Swedish; and be at least 18 years old. The general exclusion criteria were any of the following, regardless of diagnosis:

- Severe MDD combined with moderate to high risk of suicide
- Recent medication changes, and comorbid bipolar or other psychotic disorders
- Participation in concurrent psychotherapy
- Current alcohol or illicit drug abuse/dependence
- Communication difficulties that would impact treatment

The studies were approved by the Regional Ethics Board in Stockholm, Sweden. All participants provided written informed consent.

2.1.2 Study sample from Uppsala University (Study II)

Participants diagnosed with SAD based on MINI were randomly selected for ICBT or placed on a waitlist control group that received treatment after 9 weeks. Exclusion criteria were the same as those described in section 2.1.1. The Regional Ethical Review Board in Uppsala, Sweden, approved the study, and written informed consent was obtained from all participants.
2.1.3 Study sample from UK and Germany (Study IV)

Participants with PD (35%), panic disorder with agoraphobia (PD/AG = 41%), or specific phobia (SP = 19%) were included. Diagnoses were made according to DSM-IV criteria using MINI,\textsuperscript{89} the Diagnostisches Interview bei Psychischen Störungen (DIPS) or the Composite International Diagnostic Interview (CIDI). Participants were enrolled at the Mental Health Research and Treatment Centre, Ruhr-Universität, Bochum and the Dental Clinical Bochum, or at the Technische Universität in Braunschweig, Germany. Participants with PD/AG were enrolled from two multi-center randomized controlled trials of CBT. The exclusion criteria were the same as those described in section 2.1.1. All trials were approved by the human ethics committees of the relevant institutions, and all participants provided written informed consent.

2.2 THE INTERVENTION

2.2.1 ICBT (Study I–IV)

The version of ICBT used in Studies I–IV were therapist guided. Essentially, it could be described as bibliotherapy administered via a secure online platform with continuous support from a psychologist. Evidence-based manuals were used and packaged as modules consecutively presented on a weekly basis for the patient. The different treatment programs for each particular disorder varied in length, but typically ranged from 9–15 modules. The patients were estimated to complete approximately one module per week, and were granted access to the following module on completion of the previous one. Participants had access to a secure online mailing system in which they could communicate with their psychologist. For an outline of the ICBT procedure, see Figure 4.

![Figure 4](image-url)
2.2.1.1 ICBT for SAD

The manual for the ICBT program was taken from a cognitive behavioral model of SAD based on the work of Clark and Wells (1995),\textsuperscript{108} their model was later adapted for ICBT by Hedman et al. (2011).\textsuperscript{105} In this therapeutic model of SAD, one of the main objectives was to stop the cycle of fear avoidance. This was done by addressing the various aspects of anxiety, catastrophic interpretations, the focus on internal and external threats, and by mapping safety and avoidance behaviors. Weekly home assignments were used to ensure continuous training by the patient, with tasks typically focusing on cognitive restructuring and gradual exposure to social behavior. For an overview of the components of the treatment program, see Table 1.

**Table 1.** Content of Internet-delivered cognitive behavior therapy for social anxiety disorder.

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module 1: Psychoeducation about social anxiety</td>
<td>Characteristics of social and general anxiety and fear; impact of cognitions and behaviors on fear and avoidance.</td>
</tr>
<tr>
<td>Module 2–3: Register social fear and avoidance</td>
<td>Mapping avoidance behavior and safety behaviors.</td>
</tr>
<tr>
<td>Module 4–10: Gradual social exposure</td>
<td>Gradually exposing oneself to social and other anxiety-provoking situations.</td>
</tr>
<tr>
<td>Module 6–11: Behavioral experiments</td>
<td>Planning and performing behavioral experiments with detailed predictions and validations of the situation.</td>
</tr>
<tr>
<td>Module 12: Relapse prevention</td>
<td>Summarizing treatment; preparing for early detection of future signs of relapse.</td>
</tr>
</tbody>
</table>

Note: This was a 14-week program with 9–15 modules, depending on the clinical site.
2.2.1.2 ICBT for PD

The ICBT manual used for PD was based on the model by Clark et al. (1994); it was later revised and adapted by Bergström et al. (2009). The treatment program for PD included psychoeducational learning of physical fear responses, cognitive restructuring, exposure training, and relapse prevention distributed across 10 modules. See Table 2 for an overview of the modules included in ICBT for PD.

Table 2. Content of Internet-delivered cognitive behavior therapy for panic disorder.

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module 1: Psychoeducation about anxiety</td>
<td>Characteristics of anxiety and fear; impact of cognitions and behaviors on fear and avoidance.</td>
</tr>
<tr>
<td>Module 2–3: Register fear and avoidance</td>
<td>Mapping avoidance behavior and safety behaviors; building anxiety hierarchies.</td>
</tr>
<tr>
<td>Module 4–6: Interoceptive exposure training</td>
<td>Gradual exposure exercises to provoke internal/bodily triggers (e.g., elevated heart rate, dizziness, etc.).</td>
</tr>
<tr>
<td>Module 6–9: In vivo exposure training</td>
<td>Gradual exposure exercises to fear-provoking situations in real life.</td>
</tr>
<tr>
<td>Module 10: Relapse prevention</td>
<td>Summarizing treatment; preparing for early detection of future signs of relapse.</td>
</tr>
</tbody>
</table>

Note: This was a 12-week program with 10 modules.
2.2.1.3 ICBT for MDD

Just like in ICBT for SAD and PD, the content of the ICBT program for MDD was the same as that of traditional CBT. The CBT protocols for MDD were developed by Martell (2001)\textsuperscript{111} and Beck (1979)\textsuperscript{112}, and include modules for cognitive restructuring, behavioral activation, and relapse prevention (see Table 3 for a description of the ICBT content for MDD). The core component of ICBT for MDD, which also differed from the other treatments in this thesis, was the behavioral activation module. Behavioral activation is the process in which emphasized attempts were made to expand the overt behavioral repertoire to increase reinforcing environmental contingencies. In addition, scheduled activities were planned to discover avoidance behaviors and other cognitive processes (e.g., ruminating) involved in the shortfall of reinforcing activities.

**Table 3.** Content of Internet-delivered cognitive behavior therapy for major depressive disorder.

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module 1: Psychoeducation about depression</td>
<td>Characteristics of depression; impact of cognitions and behaviors on depression.</td>
</tr>
<tr>
<td>Module 2–6: Behavioral activation</td>
<td>Increasing positively reinforced activities (e.g., physical activity, social encounters, etc.), decreasing avoidance behaviors (e.g., binge eating, excessively watching tv), and handling short-term negatively reinforced behaviors (e.g., paying bills, cleaning, etc.).</td>
</tr>
<tr>
<td>Module 4–9: Cognitive reappraisal</td>
<td>Recognizing negative thought patterns, and registering and validating them through behavioral experiments.</td>
</tr>
<tr>
<td>Module 10: Relapse prevention</td>
<td>Summarizing treatment; preparing for early detection of future signs of relapse.</td>
</tr>
</tbody>
</table>

Note: This was a 12-week program with 10 modules.

2.2.2 Group-delivered CBT (GCBT) (Study I and IV)

These interventions were based on standard CBT treatments for SAD and PD, and the modules were similar to the content of the ICBT for SAD and PD (see Table 1 and Table 2). In GCBT for SAD, treatment consisted of one initial individual session followed by
fourteen weekly group sessions over 15 weeks. In GCBT for PD, patients were given ten sessions over the course of 12 weeks. Both groups were led by two clinical psychologists.

2.2.3 Standard CBT (Study IV)
In the Ruhr-University and Braunschweig cohorts, participants received standard exposure-based CBT for SP, AG, or PD. The specific treatment details are described in detail elsewhere.\textsuperscript{113} The subgroup of patients with SP (dental patients), received a shorter, dental phobia-specific exposure treatment program.

2.3 PRIMARY OUTCOME MEASURES

2.3.1 Liebowitz Social Anxiety Scale (LSAS)
The primary measure of the outcome for social anxiety treatment in Study I and II was the clinician-rated LSAS.\textsuperscript{114} The psychometrics for both scales are considered reliable, validated, and sensitive to SAD treatment.\textsuperscript{115} The LSAS measures fear and avoidance on a Likert-type scale ranging from 0 (no fear/never avoid) to 3 (severe fear/usually avoid) for either performance situations or social interaction situations. In both trials, participants provided data at pre- and post-treatment, and at follow-up.

2.3.2 Self-rated Montgomery Åsberg Depression Rating Scale (MADRS-S)
The primary outcome measure for Study III was MADRS-S. The total MADRS-S score, which ranges from 0 to 54, measures nine clinical characteristics of depression, and has good psychometric properties.\textsuperscript{116,117} MADRS-S baseline (MADRS-S Pre) was assessed at treatment start, once per week during treatment, and at the end of treatment (MADRS-S Post). Thus, each individual provided up to twelve weekly MADRS-S assessments that were included in the analyses.

2.3.3 Panic Disorder Severity Scale (PDSS)
The primary measure of the outcome for the Swedish sample in Study I, and in one of the four subsamples (Karolinska cohort) of Study IV, was a self-rated PDSS (PDSS-SR).\textsuperscript{118} The PDSS-SR is validated and shows moderate to high psychometric characteristics. The PDSS-SR has seven items, each with a 5-point scale ranging from 0–4, giving a total score range of 0–28. The scale assesses the frequency and severity of panic attacks, anticipatory anxiety, phobic avoidance, and occupational and social impairment.
2.3.4 Clinical Global Impression-Severity (CGI-S)
In one of the subsamples of Study IV, the primary outcome measure used to assess symptom severity was the validated clinician-rated CGI-S. The CGI-S ranges from 1 to 7, where a score of 1 indicates that the patient is healthy and a score of 7 is indicative of severe illness. The CGI-S was rescaled to a range of 0–6, so that full remission could be represented as 100% change.

2.3.5 Panic Agoraphobia Scale (PAS)
In a subsample of the Panic-Net Consortium in Study IV, the primary outcome measure was the validated self-rated PAS. The scale includes 14 items; the first item is a screening item, and the other 13 are used to determine symptom severity. Each item has a scale of 0–4, with a maximum total score of 52. The scale measures the frequency, severity, and duration of panic attacks; agoraphobic avoidance; anticipatory anxiety; impairment; and worries about health.

2.4 TREATMENT RESPONSE

2.4.1 Study I and II
Treatment response was measured by the end-state symptom severity reported through LSAS and LSAS-SR, the status of diagnosis after treatment, and the criteria for clinically significant improvement—also known as the reliable change score, defined by Jacobson and Truax (1991). The reliable change score takes both statistical and clinical significance into account by stating that participants had to show improvement by being closer in terms of symptoms to the healthy population than to the SAD population.

2.4.2 Study III
Treatment response was measured as a consistent significant effect on the MADRS-S scores or changes in the MADRS-S scores over time.
2.4.3 Study IV
Treatment response was calculated by examining the changes in symptom severity from start-of-treatment (pre) to end-of-treatment (post). Scores were converted to a percentage of the maximum possible score for each scale for comparison across sites.

2.5 COLLECTING GENOTYPES

2.5.1 DNA collection
After informed consent was obtained, all participants from the Swedish samples provided a blood sample. The blood was stored in a freezer locally at the hospital or sent to the Karolinska Institutet biobank for DNA extraction. DNA from subsamples in Study I and Study IV were obtained from saliva samples.

2.6 STATISTICAL ANALYSIS

2.6.1 Study I and II
In Study I, statistical analyses were performed using PASW version 18.0 (SPSS, Chicago, IL, USA) and ROC4 (Stanford University, Stanford, CA, USA); in Study I, SPSS 20.0.0 (SPSS, Chicago, IL, USA) and Statview version 5.0. (SAS, Cary, NC, USA) were used. We used regression analysis to identify predictors of treatment outcome. Each regression model contained LSAS baseline values, the potential predictor variable, the treatment condition/site, and the interaction factor. The effects of genotype on treatment response were analyzed by repeated measures analysis of variance (ANOVA), and linear trend analysis was performed based on weekly measure points in Study II.

2.6.2 Study III
In every analysis, we used all available data points for all patients and all statistical analyses were performed using R software. We used full interaction maximum likelihood mixed models to analyze the association between the GRS values and ICBT treatment outcomes measured by MADRS-S. We fitted a regression model that included the linear and quadratic effects of time (to allow for curvilinear development over time, which provided the best fit of the data). Next, we investigated the influence of GRS on the rate of change during treatment. In all models, covariates (e.g., GRS) and possible confounders
(e.g., principle components ancestry [PCA], age, etc.) were added as both main effects and interaction effects with the linear and quadratic effects of time. A significant main effect means having a constant effect on MADRS-S over the treatment time. The interpretation of a significant GRS × time interaction effect was if the GRS influenced the rate of improvement during treatment. We examined the utility of GRS for predicting the degree of improvement from pre- to post-treatment. We tested for an interaction between MDD GRS and treatment response (drop in MADRS-S score × random effect of time).

2.6.3 Study IV
Given that the subsamples in Study IV differed in their primary outcome measures, a mixed linear model association was used. Next, analyses were performed with a linear regression model of the dependent percentage change outcome and independent covariates (baseline severity, site, number of modules, age, etc.) Missing data was handled using the last observation carried forward, and was later included as a covariate in the regression analysis. The association statistics from the primary analysis of the adult sample were subjected to a meta-analysis with the results from the treatment-response GWA study of children with anxiety disorder. The common genetic variant heritability of the meta-analysis sample was then assessed by performing linkage disequilibrium (LD) score regression.

2.7 GENOTYPING
2.7.1 Study I
DNA extraction from whole blood was performed using standard protocol. For the bi-allelic 5-HTTLPR sequence, two fragments of 336 bp (S) and 379 bp (L) in length were amplified by PCR and separated by agarose gel electrophoresis. The rs25531 SNP was identified as previously described. To genotype COMT V158M (rs4680) and BDNF V66M (rs6265), we used the TaqMan allelic discrimination assay and the ABI instrument [HT7900] (Applied Biosystems) as previously described. All genotyping experiments were performed in duplicate.

2.7.2 Study II
DNA was extracted from whole blood, or from saliva using the Oragen Purifier (www.dnagenotek.com). COMT V158M (G472A, rs4680) and TPH2 G703T (rs4570625) were genotyped with TaqMan SNP genotyping assays and an ABI 7900 HT instrument
(Applied Biosystems) under standard conditions. The two fragments of the bi-allelic 5-HTTLPR, the 336-bp (S) and the 379-bp (L) fragments, were amplified by PCR and separated by agarose gel electrophoresis. All genotypes were dichotomized, the samples were genotyped in duplicate, and genotype assessors were blinded to case-control status.

2.7.3 Study III
Genotyping was performed at Life & Brain GmbH (Bonn, Germany) using the Infinium Global Screening Array 1.0 BeadArray (Illumina, San Diego, CA, USA) and automated workflow according to the manufacturer’s instructions. Raw data were analyzed using GenomeStudio 2.0 (Illumina) using the Infinium cluster file (GSA-24v1-0_A1_ClusterFile.egt). For re-clustering, the GenTrain 3 algorithm in GenomeStudio 2.0 was used.

2.7.4 Study IV
Genotyping of the Bochum participants was performed using the Illumina PsychChip microarray (Illumina) at the Institute of Psychiatry, Psychology and Neuroscience, King’s College London, UK. The Swedish subsample were genotyped using Illumina HumanOmniExpress BeadChips (Illumina) at the Department of Genomics, Life & Brain Centre, University of Bonn, Germany. The Panic-net samples were genotyped using Illumina Human660W-Quad BeadChips and Sentrix BeadChip Array HumanHap300 Genotyping BeadChips (Illumina).

2.7.5 Quality control
Quality control using PLINK 1.9 software\textsuperscript{124} was performed for each cohort in Study III and IV as previously described. The GWA data from Study III were processed using the Psychiatric Genomics Consortium (PGC) Ricopili pipeline for quality control and genotype imputation with reference genomes from the 1000 Genomes Project (phase 1 version 3).\textsuperscript{125} In Study IV the quality controlled data were phased using SHAPEIT.\textsuperscript{126} Variants were excluded if they were rare (minor allele frequency < 0.05) or not present in > 99% of participants in Study II and IV. Participants in all studies were removed if they deviated substantially from Hardy-Weinberg equilibrium (HWE; $p < 1 \times 10^{-5}$). Every HWE model was tested using a chi-square test. Participants with hidden relatedness or sample duplicates
were removed; for all pairs with PI-HAT > 0.2, one member of each related pair was removed.

2.7.6 Principal components ancestry (PCA)
We examined the population substructure by principal component (PC) estimation. Principal component analysis is a data reduction strategy; in genotyping quality control, it helps to reduce the amount potential confounders due to population stratification. Population stratification occurs when there are systematic ancestral differences between the allele frequencies in the control sample. In Study III, we tested twenty PCs for phenotype association (using logistic regression with appropriate covariates) and evaluated their impact on the genome-wide test statistics using lambda GC (expected value is 1.0 in the case of no inflation) after GWA analysis of the specified PCA. Forty-nine samples were excluded from the analysis as outliers in both HapMap3 and 1KG. In Study IV, PCA analysis was performed using EIGENSOFT. Outliers were removed if they were > 6 standard deviations from the mean of the first three PCAs.

2.7.7 Polygenic risk scores (PRS)

2.7.7.1 Study III
PRSs were generated for the following six phenotypes: MDD, BIP, ADHD, autism spectrum disorder (ASD), intelligence quotient (IQ), and educational attainment (EDU). We obtained the corresponding GWA study results for MDD, BIP, ADHD, and ASD from the PGC website (https://www.med.unc.edu/pgc/results-and-downloads), and the GWA study results for IQ and EDU from published GWA meta-analyses. The PRSs were categorized into tertiles. For each subject in the target dataset, the MDD PRS was calculated using the PLINK “score” procedure.

2.7.7.2 Study IV
Summary statistics from GWA studies in PRSice were used to generate PRSs for all participants in the adult cohort. The PRSs were generated for the same phenotypes as in Study III, with the addition of SCZ but without the ASD. PRSice performs multiple regressions, incorporating differing numbers of genetic variants into the polygenic scores.
determined by discovery trait GWA studies with eight p-value thresholds: \( p = 1, p = 0.5, p = 0.1, p = 0.05, p = 0.01, p = 0.001, p = 5 \times 10^{-6}, \) and \( p = 5 \times 10^{-8}. \)

2.8 ETHICAL CONSIDERATIONS
All studies were conducted in accordance with the principles outlined in the Declaration of Helsinki. Ethical approval and informed consent were obtained from all participants, and rules and regulations regarding biobanks met all the criteria for all studies in this thesis.
3 EMPIRICAL STUDIES

3.1 STUDY I

Aim:
The aim was to identify clinical and genetic predictors for patients with SAD and participant response to CBT.

Method:
Participants were recruited from a randomized controlled trial of adult patients (n = 126) who received either GCBT or ICBT for SAD. We investigated potential demographic (e.g., age), clinical (e.g., comorbid depression), and treatment-related predictors (e.g., adherence), and genetic predictors. Data on non-genetic predictors were collected in clinical interviews and using self-report measures. Participants were genotyped from whole blood on three candidate gene variants involved in dopaminergic and serotonergic processes: 5-HTTLPR and a SNP in close proximity (rs25531), COMT V158M, and BDNF V66M.

Three types of statistical analyses were performed: i) linear regression was used to analyze social anxiety (for each variable, a regression model was built using LSAS scores as a dependent variable), ii) logistic regression using SAD as the dependent variable, and iii) signal detection analysis that outputs the receiver operator characteristics (ROC) of subgroups with high and low predictor cut-offs.

Results:
Neither of the genetic polymorphisms were associated with treatment response in this study. The significant clinical predictors was; having children, full-time employment, less severe depressive symptoms, adherence to treatment, and higher expectations for CBT, and taken together they explained about 50% of the variance in treatment response.

3.2 STUDY II

Aim:
The aim of Study II was to investigate whether three genetic polymorphisms predicted CBT responses to SAD in a larger sample than previously investigated (in Study I).

Method:
Participants were recruited from two separate RCTs of CBT for SAD at two independent sites in Sweden. In trial 1 (n = 112), participants were assessed before treatment,
weekly, after treatment, and at 6 months follow-up. In trial II (n = 202), participants were assessed at the same time points, except for the weekly measures, and the follow-up was performed after 12 months. The primary outcome measure was the LSAS-SR, and improvement over time was associated with polymorphic variations in three genetic variants; the S carriers of 5-HTTLPR, the M/M genotype of COMT V158M, and the T carriers of TPH2 G703T. Linear regression models were used to test the association between predictors and treatment outcome. Covariates were tested with an unpaired Student’s t-test for normally distributed variables, and the Mann-Whitney U-test was used for non-parametric variables. Association analysis of genotype and treatment outcome was tested with repeated measure ANOVA, and a linear trend analysis for inclusion of all data points. Significance levels were set at $p < 0.05$. DNA was extracted from whole blood in trial I and from saliva in trial II. All genotype frequencies conformed to HWE, as confirmed by a chi-square test. All genotypes were categorized into: S carrier (SS, SL)/non-carrier (LL) for 5-HTTLPR, V carrier (VV, VM)/non-carrier (MM) for COMT V158M, or T carrier (TT, GT)/non-carrier (GG) for the promoter SNP (rs4570625). All analyses were performed separately in each trial and pooled together.

**Results:**

In trial I, a significant interaction effect between the TPH2 G703T polymorphism and LSAS-SR was observed between the baseline and post treatment. The T carriers showed significantly better improvement than the non-carriers. This effect was only present directly after treatment, and not at 6 months follow-up. Furthermore, neither of the two other polymorphisms (5-HTTLPR and COMT V158M) showed any significant main or interaction effects at any of the given timepoints. In trial II, a significant linear trend interaction result was also demonstrated for the TPH2 G703T polymorphism, meaning that one genotype predicted greater improvement over the treatment period. However, this improvement was reversed in comparison to the result in trial I; the non-carriers (GG) performed better in this trial. The other polymorphisms (5-HTTLPR and COMT V158M) yielded no significant findings, consistent with the results in trial I. When the data from trial I and II were pooled, analysis of the three polymorphisms failed to detect any associations between genotype and CBT response.
3.3 STUDY III

**Aim:** The aim of Study III was to generate polygenic risk scores from six psychiatric and cognitive traits and investigate their predictive power for ICBT response in MDD in 894 adult patients.

**Method:**
Participants were recruited between 2008 and 2016 from regular patients receiving care at the Internet psychiatry clinic in Stockholm, Sweden. The clinic is a government-funded health care unit within a regular psychiatric setting. The participants had either been referred to the clinic by their general practitioner or via an online self-referral system. All participants fulfilled the criteria for MDD and received ICBT. The primary outcome measure was MADRS-S, and was assessed before treatment, weekly, and after treatment.

PRS were calculated from discovery data sets for six phenotypes: MDD, BIP, ADHD, ASD, IQ, and EDU. Participants were excluded from the target set based on genotyping-related issues (sample overlap, cryptic relatedness, poor call rate, and non-European ancestry [population stratification]), resulting in a final target sample of 894 participants. The GRS sums were calculated from the scores based on the risk alleles weighted by the effect size in the discovery sample. To generate GRSs in an independent sample, we performed LD clumping ($r^2 < 0.1$ in a 1-Mb window). For LD reference samples, we used overlapping SNPs from European samples of the 1000 Genomes Project. We calculated the GRS value for each phenotype using a cutoff value of $p \leq 0.05$. For statistical analyses of full interactions, maximum likelihood mixed models were used. These models allowed us to include all available data from all patients. Covariates (e.g., sex, age, and PCA) were included in the model and outlier analysis was performed on cases that could disrupt the regression model. However, removing potential outliers from the model did not affect the interpretation of either the main or the interaction effects.

**Result:**
The main finding in our primary analyses was a significant interaction effect ($B = 0.09$, $p < 0.001$) for the ASD GRS and MADRS-S changes over the treatment period (Figure 5). This means that a higher ASD genetic load was correlated with a reduced decline in MADRS-S over time. Further, as expected, we observed a general negative treatment effect over time ($B = -1.29$, $p < 2 \times 10^{-16}$) and a significant effect of quadratic time on MADRS-S ($B = 0.048$, $p < 2 \times 10^{-16}$), regardless of genotype. This was understood as the general effect of the treatment, regardless of genotype, meaning that the MADRS-S scores decreased during treatment, with a larger decline in the beginning. None of the other phenotypes (including GRS MDD) had a robust effect on MADRS-S or on MADRS-S over time.
Figure 5. Effect of ASD GRS at $p$-value threshold 0.05 on MADRS-S scores during CBT for MDD. The figure shows the predicted MADRS-S score for every week during treatment for three different levels of the ASD GRS (25th, 50th and 75th percentiles). The shaded area shows the 95% confidence intervals of the predicted values. Participants with the highest ASD GRS scores (blue) showed poorer responses to treatment versus those with average (green) or low (red) ASD GRS scores. Abbrevations: autism spectrum disorder (ASD), genetic risk score (GRS), Montgomery Åsberg Depression rating scale -self (MADRS-S), cognitive behavior therapy (CBT).
Note: The Figure and figure legend are included in Study III.

3.4 STUDY IV

Aim:
The aim of Study IV was to investigate the genetic effects underlying CBT outcome, and to study the genetic overlap in psychiatric and cognitive traits through a GWA meta-analysis and by calculating PRS for psychiatric and cognitive traits.

Method:
We performed a GWA study of the response to CBT in adults with anxiety disorders ($n = 972$),\textsuperscript{135-138} followed by a meta-analysis of our results with a comparable sample of
children\(^{139}\) (child sample \(n = 939\); meta-analysis \(n = 1911\)). We estimated the variance in treatment response that could be explained by common genetic variants (SNP heritability), and polygenic scoring was used to examine the genetic overlap of treatment response with psychiatric disorders and learning abilities. Participants were drawn from one of three broad studies of CBT. Diagnoses were made according to DSM-IV criteria using the MINI 5.0 or the DIPS, or the CIDI. The three predominant disorders were PD (37%), PD/AG (42%), and SP (19%). These disorders share the common components of excessive fear, anxiety, and avoidance behaviors. All participants (\(n = 972\), 60% female) received CBT for an anxiety disorder. The child sample consisted of 939 children and adolescents (aged 5–17) with DSM-IV criteria indicating anxiety disorder diagnoses, who received CBT at one of eleven sites. Outcome analyses examined the percentage change of symptom severity from start-of-treatment (baseline) to end-of-treatment (post-treatment). As continuous outcome measures differed between cohorts, they were standardized. In brief, DNA from 966 participants was extracted from blood by routine desalting methods. For six Bochum participants and twenty-eight Braunschweig participants (3.5% of total sample), DNA was obtained from saliva samples. The effects of clinical covariates on symptom severity at baseline and post-treatment, as well as treatment outcomes, were assessed using linear mixed models.

**Result:**

Phenotypes and high-quality genotype data were available for 972 individuals. After quality control, imputation, and genetic data-merging procedures, a total of 4.74 million genetic variants were included in the analysis. No individual genetic variant was statistically significant after correction for multiple testing \((p < 5 \times 10^{-8})\). However, eight independent genetic loci surpassed the threshold for potential significance \((p < 1 \times 10^{-5})\). Eight independent genetic loci were identified on chromosomes 3, 6, 7, 8, 9, and 16 in the adult sample (see Figure 6), and a further eight loci were identified on chromosomes 3, 4, 7, 9, 10, and 20 in the meta-analyzed sample (see Figure 7). The estimate for SNP heritability was moderate, but non-significant \((\text{SNP } H^2 = 0.24, \text{SE} = 0.31)\). None of the polygenic scores tested, including EDU, were significantly associated with the response to psychological therapy after applying statistical corrections.
Figure 6. A Manhattan plot of the p-values from genome-wide associations with response to psychological therapy in the adult sample (n=972). The x-axis displays associated genetic variants, arranged by location on the chromosome. The y-axis shows the strength of the association with treatment response. The red line represents the conventional threshold for genome-wide significance ($p=5\times10^{-8}$) and the blue line represents a conventional threshold for suggestive significance ($p=1\times10^{-5}$).

Note: The Figure and figure legend are included in study IV.
Figure 7. A Manhattan plot of the p-values for genetic associations with response to psychological therapy from the meta-analysis of the adult (n=972) and child samples (n=939; total n=1,911). The x-axis displays associated genetic variants, arranged by location on the chromosome. The y-axis shows the strength of the association with treatment response. The red line represents the conventional threshold for genome-wide significance ($p=5\times10^{-8}$) and the blue line represents a conventional threshold for suggestive significance ($p=1\times10^{-5}$). Note: The Figure and figure legend are included in study IV.
4 GENERAL DISCUSSION

4.1 CAN CLINICAL VARIABLES AND CANDIDATE GENE POLYMORPHISMS PREDICT CBT TREATMENT RESPONSE IN SAD?

Study I was the first study to look at both clinical and genetic variables as potential predictors for CBT response. The significant clinical predictors in Study I, explained half of the variance in CBT treatment outcome, indicating essential predictive information to guide clinicians in the process of delivering CBT both in groups and via the internet. The results are in line with a previous study of patients in a routine psychiatric care setting showing an association of better MDD CBT response and full-time employment. In both studies the treatment was strictly manualized and monitored, but the disorders differed (SAD and MDD), suggesting that treatment format rather than diagnosis might be of relevance. Perhaps the heterogeneity in treatment settings, format and how treatment outcome are measured, explains some of the contradictive results in studies of predictors and psychological treatment outcome. However, given the sparse evidence in the literature regarding predictors of CBT outcome, it is important to consider these findings as preliminary until proper replication can be performed.

The main novelty in Study I was the investigation of three common genetic variants and their possible predictive value for treatment outcome. None of the genetic variants were associated with treatment outcome in this study; thus, we were unable to replicate the previous study on COMT V158M and CBT outcome. However, in that study, the patients received CBT for PD, and the therapy consisted of gradual exposure to feelings of panic and extreme acute fear. In addition, the same study only found an effect relating to the exposure part of the treatment; when including the treatment as a whole, including psychoeducation and mapping avoidance and safety behaviors (see Table 4 for a full overview), the effect of the polymorphism disappeared. Variants in the genes encoding 5-HTT, COMT, and BDNF have all been linked to processes relevant to treatment outcome. However, given the many negative findings in candidate previous gene studies, this null finding is not surprising given that the contribution of a single gene variant is now thought to be negligible in explaining the variation in a complex behavior such as treatment response. Still, it is noteworthy that one of two genetic loci detected in the largest ever GWA meta-analysis of anxiety (PD, not SAD), showed an association at a locus in the COMT gene. Perhaps more and bigger studies into the relationship between the COMT pathway and anxiety, as well as subsequent responses to anxiety treatment, might provide some insight to the molecular processes involved. A consideration in this study is the inherent limitation of recruitment from an RCT. The
predictive power of any variable was most often small, and some of the predictors expected to have the highest value (e.g., suicidal thoughts) are commonly excluded from an RCT. Nevertheless, this study allowed for comorbidity and psychotropic medications; thus, it had rather tolerant restriction criteria and should be considered to be fairly ecologically valid.

4.2 CAN WE DETECT GENETIC PREDICTORS OF CBT RESPONSE IN SAD IF WE DOUBLE THE SAMPLE SIZE THROUGH A MULTI-CENTER APPROACH?

The aim of this study was to investigate the candidate gene polymorphisms and their relationships with CBT response in SAD in a sample from two clinical sites. The gene variations in 5HTT/SLC6A4, TPH2, and BDNF were investigated. The findings in the Stockholm sample and Uppsala sample showed significant associations with CBT response; however, the genotypes showed an opposite direction for different trials. In the Stockholm sample, carriers of the T allele of the TPH2 G703T mutation showed a significant interaction effect with time and decreased symptom scores directly after treatment; conversely, in the Uppsala sample, the significant effect of the homozygote GG genotype predicted a more positive response after 6 months. In an earlier study of amygdala responsiveness and the TPH2 T carrier genotype in SAD, an effect for the genotype was found; however, this effect was not found in cases, only in healthy T carriers. These two studies (with three samples), which are currently the only studies investigating the association between TPH2 G703T and SAD, yielded significant results, but in three different directions. Yet, the variety of strategies to measure treatment response in the different studies varies from symptom decrease to experimental fear extinction protocols, which might affect the power to detect genetic signals. In addition, in this study the phenotype (CBT response) was observed within an RCT setting. This generally means that the therapists commonly are especially well-trained and may have been more motivated to assist the patients. Additionally, the level of monitoring and control of the patient is usually higher for an RCT, reducing the proportion of patients who would drop out, thus increasing the likelihood of a positive CBT response. The variance in the RCT population might be smaller than that in a routine psychiatric care setting, implying that the variance to detect genetic effects would decrease.
4.2.1 Methodological considerations in Study I and II

In both Study I and Study II, the participants were recruited from an RCT, resulting in the exclusion of some potential predictors (e.g., suicidal thoughts, etc.). It has been suggested that there may be a link between a larger genetic load and a more severe or disabling phenotype; with that in mind, excluding the most severe or burdened cases might have significantly reduced the predicting power. In both studies, we decided to classify the genotypes in bi-allelic categories to maximize predictive power and to follow previously published studies. We divided the alleles into classes of potential “risk allele carriers” (S carrier, M carrier, and T carrier) and “non-carriers,” instead of the following threefold classes: homozygous risk carrier (e.g., SS), heterozygote (e.g., SL), and homozygote non-carrier (e.g., LL). Nonetheless, if we were to find a genetic predictor in that sample, the probability of that finding being a false-positive would have been large because of the missing power to detect predictors with small effect sizes. Based on current knowledge, the largest limitation in both Study I and Study II, given the sample size, was the use of the candidate gene approach to detect genetic signals for CBT response within a study with our sample size. In both studies, the power to detect a genetic predictor from a few candidate genetic variants was very low. One reflection regarding all candidate gene approach studies is that the standard probability value of \( p \leq 0.05 \) would be too facile, and would result in many false-positive findings. Even when the sample size was doubled between Study I and Study II, sample size was found to be only one part of the issue of predictive power. To detect any possible genetic signals, substantially larger samples are needed; most of all, a different strategy is required that complements the idea of single critical genes into genome-wide methods without any \textit{a priori} hypothesis on the genetic pathways involved. In sum, we can conclude that there was reason to investigate the genetic variants in our earlier candidate gene studies, given the previous findings, but they were underpowered to detect robust findings.

4.3 CAN AGGREGATED GENETIC SCORES BASED ON GENOME-WIDE ASSOCIATION STUDIES ON PSYCHIATRY-RELATED TRAITS PREDICT TREATMENT RESPONSE IN MDD?

The aim of Study III was to identify genetic variants in relation to CBT responses in MDD using a genome-wide approach where variation profiles associated with six other traits were investigated by calculating polygenic risk scores. This was enabled through a two-step process: first, by conducting a GWA study of response to CBT in MDD responses; and
second, by calculating PRSs to study the utility of these genetic variants in predicting CBT treatment outcome in MDD. None of the results in the GWA analysis were significant; this result was expected, given the large sample sizes required for other complex traits (e.g., MDD required approximately 130,000 subjects). However, when we in the second step calculated the six PRSs based on associations with other phenotypes and correlated them with depression symptoms over time, one was significant: the PRS based on the ASD-associated gene variation profile. This means that a patient with a higher genetic load for ASD, also has an increased risk of not responding as well to treatment compared to those with a lower genetic ASD load. Inferring that a high ASD GRS also implies a more highly expressed phenotype of ASD, we can speculate as to why these individuals could be at risk for a worse CBT treatment response. In Study I, participants with a higher adherence and expectations for the treatment responded more effectively to CBT. If a person with both MDD and ASD-related difficulties were to read and learn about negative current thoughts and a deficit in positively reinforcing behaviors (the core interventions in CBT for MDD), they might have low expectations of the treatment due to an inadequate or insufficient description of the disorder and its treatment. This would be in line with predictor studies of CBT response and the relevance of a high treatment expectancy and adherence for a positive outcome. If a person with ASD-related problems, such as difficulties with social interaction and communication, only learns about the psychoeducational aspects of MDD, their expectations and adherence might be lower. A limitation to this explanation is that very few of the subjects in this study had an ASD diagnosis. In a future study, it would be interesting to further explore and establish the relationship between ASD GRS and ASD expressed phenotypes. Another noteworthy result of our study was the fact that MDD GRS was not associated with MDD CBT outcome. This might indicate that the genetic background for MDD differs from that of treatment effectiveness. There has been a tentative hypothesis in many studies of CBT response that the predictor (e.g., genetics, neurobiological pathways, etc.) of the disease also plays a part in the response to treatment. This preliminary finding raises the question of whether psychological treatment response involves an entirely different biological process, or maybe just partially, or only under certain circumstances.

4.3.1 Methodological considerations for Study III
Since the therapygenetics field expanded its focus from only investigating a few critical genes, to studying the aggregated genetic loads in complex traits, this study was the first to find a result on psychological treatment response and genetics. However, this result needs to
be considered preliminary. A previously described phenomenon is the tendency for the first study in a field to overestimate the effect size of its findings, and subsequent studies to moderate the initial effect. Given that it took several years of major effort to collect a large enough sample size to find a robust genetic signal for MDD, there is likely that future studies into treatment response will need even larger samples to find robust and replicable findings.

4.4 CAN WE ESTIMATE GENETIC VARIANCE IN TREATMENT RESPONSE FROM A GENOME-WIDE STUDY IN ANXIETY DISORDERS? CAN WE IDENTIFY GENETIC OVERLAP OF CBT RESPONSE AND PHENOTYPES FOR PSYCHIATRIC AND COGNITIVE TRAITS?

The fourth and final study was a genome-wide association study on CBT treatment response in anxiety disorders. Summary statistics from the adult sample analysis were then subjected to meta-analysis with the summary statistics from the child GWA study of the response to CBT (total n = 1911). No genetic variants were associated with treatment response after correction for multiple testing. However, eight independent genetic loci on chromosomes 3, 4, 7, 9, 10, and 20 reached near-significance ($p < 1 \times 10^{-5}$) in the adult sample, along with another eight loci on chromosomes 3, 4, 7, 9, 10, and 20 in the meta-analysis sample. Of the putative loci for treatment response, genotypes involved in psychiatric phenotypes and learning abilities have been suggested. In line with the preliminary findings of Study III, two of the suggested loci hits have previously been linked to ASD; the suggestive top hit on chromosome 6, MIR1234 and rare mutations in the DLGAP2 gene on chromosome 8, have been associated with ASD in two studies. In the other suggestive hits, variants associated with for instance; reading ability, SCZ, loneliness, ADHD, brain volume in infants, and subjective well-being were identified; all of these are traits (psychiatric and cognitive) that can be assumed to play a part in treatment response behavior.

In this study, none of the investigated polygenic scores (anxiety, MDD, SCZ, EDU and IQ) were associated with CBT response. No autism-related PRS was calculated, so a replication of the preliminary findings in Study III was not possible. We expect that the genetic effects that underpin treatment outcome will be shared with psychiatric disorders and related cognitive traits in future high powered studies.
4.4.1 Methodological considerations for Study IV

There are different aspects to consider in understanding the results in Study IV. First, the baseline severity for all participants was controlled for, meaning that genetic signals could only be detected beyond baseline severity. It is promising that the prospects for translating polygenic risk into clinical use are increasingly investigated.\(^{156}\) Secondly, regarding the power of genetic studies in predicting CBT response, there are some potential trade-offs to take into consideration. In this study, we aimed to gather as many subjects as possible with comparable genotypic and phenotypic procedures. However, the heterogeneity of the subsamples may have diluted the phenotypes and thereby failed to detect genetic signals. Experience from genetic studies of other highly heterogeneous traits, such as MDD,\(^ {48}\) show us that the need for larger samples are connected to the level of complexity of the trait. In Study III, the phenotype had a more standardized treatment setting (ICBT) and the participants were all from the same cohort. This homogeneity might have been the reason why signals could be detected in this smaller sample (Study III), and not in Study IV. However, there is also a possibility that the autism-related risk score that was investigated in Study III, and not in Study IV, is the only genetically overlapping trait with treatment response. Yet, it is likely that in future studies, the traits investigated in Study IV will contribute to the explained variance of treatment response, as both cognitive and psychiatric traits are heritable.

4.5 FUTURE DIRECTIONS

From the beginning of planning the first study in this thesis (2009) until the completion of the final study (2018), a paradigm shift in our view of psychiatric therapy genomics has taken place.\(^ {42}\) Due to inconsistent findings, contradictory results and technological development, the psychiatric genetics field developed from having a partial focus on the critical-gene contribution in complex traits to also include the investigation of aggregated small effects of common variants.\(^ {157}\)

However, with this new approach, other hurdles need to be addressed and overcome. For instance, to reduce the risk of false-positive findings in the immense number of tests performed in a GWA study, the need for new statistical standards has emerged, with more stringent \(p\)-level thresholds.\(^ {42}\) In addition, to further counteract false-positives, population stratification is essential. The need for large samples in different ethnic and geographically
separated groups are important to increase the accuracy of predictions across different allele-frequency groups. Furthermore, heterogeneity across samples might be a major difficulty to overcome. In the effort to collect massive samples, heterogeneity increases. In psychological treatment outcomes, this is even more hazardous given that both the disease and treatment are subject to extremely heterogeneous conditions. All participants in this thesis received manualized CBT, but there was still a broad range of differences between the different methods (ICBT, GCBT, standard CBT). This was possibly one of the reasons why a smaller homogeneous sample in Study III yielded a genetic signal, while the larger but more heterogeneous sample in Study IV did not. Apart from a diversity of approaches in determining phenotypes, the mere diagnostic systems (DSM and ICD) are descriptive in their nature, and do not build their classifications on etiology. This means, by definition, that a psychiatric phenotype (e.g., PD) is based on the assessed and expressed symptoms, further raising the potential for heterogeneity. Still, in a recent genome-wide study, forty-four new loci were found to be associated with MDD, assessed and diagnosed in sometimes very different manners, which implies that large enough sample sizes can handle this divergency. However, until studies of psychological treatment response can be successfully performed with larger sample sizes, the endeavor to find phenotypes as similar as possible must continue. Finally, even with large sample sizes and well-phenotyped participants, the SNP contribution in psychological complex traits, such as treatment response, may be even more polygenic than typical physical traits. The polygenicity of psychiatric traits and complex behavior raises the question on how these many genes may be coherently tied together. One plausible assumption is the possibility that genetic variation at many different loci might implicate alterations in different pathways, leading to an inappropriate response or developmental irregularity. Even well-powered GWA studies might only detect a small part of the variance, and rare SNPs (with less than 1% minor allele frequency) may play a significant part in the genetic architecture of psychological treatment responses. To detect SNPs like that would require full genome sequencing and a study of the complete DNA sequence throughout the entire genome; this genotyping method is currently prohibitively expensive, but prices are expected to become lower over time.

The field of psychiatric genetics has experienced a rapid change due to technological and computational advances, and has developed a culture of large collaborations across nations and disciplines. This experience will be essential when facing the next generation of biotechnological development and scientific progress.
There are many aspects related to genetic testing that need continuous attention and may lead to ethical dilemmas. For instance: how and what do we report back to our subjects after genetic testing? Should we inform them about incidental findings with clinical significance? Along with larger sample sizes, the probability of detecting clinically significant genetic conditions (e.g., Klinefelter syndrome) increases. Moreover, if researchers inform participants that they have genotypes with clinically significant conditions or possible risk factors, how do we set the cut-off for “possible” risk and at the same time take into account the possibility of false-positive/false-negative findings? Furthermore, how can we ensure that participants who donate DNA forever remain anonymous? In addition to confidentiality, how do we guarantee that discrimination as a consequence of genetic testing never occurs? When participants are asked to donate DNA, their genetic “fingerprint” is stored and by the uniqueness of their genetic code, the participants are by definition not anonymous. Psychiatric genetics as a field is completely reliant on the public trust and the will to voluntarily participate in research. The need for larger sample sizes must not compromise the participants’ rights to personal integrity. In addition to strict standards when handling biobank data, we need to secure participants’ rights not to know about their genetic risk for traits and diseases.

The historical misconduct in dark corners of the past have proved that we need to be alert. As practitioners of psychiatric genetics research, we need to be aware that there is negatively biased public opinion regarding biobanks, as well as great uncertainty related to what a stored DNA sample can be used for in the future. Despite the fact that the vast majority of practitioners in modern clinical psychology today embrace an understanding of psychiatric disorders as a result of both inherited and acquired characteristics, some are still concerned by the focus on the biological approach. These concerns partly relate to historical events of the terrifying mistreatment of psychiatric patients based on unscientific theories like eugenics.\textsuperscript{160} Eugenics refers to a movement that aimed at improving the genetic composition of the human race. This false idea of refining the human genome by discarding “bad genes” was the start of a chain of unacceptable events, such as involuntary sterilization, restrictions for marriages, and the institutionalization of mentally ill individuals.\textsuperscript{161} This practice had devastating consequences for the victims of these programs, and ultimately ended up being used as the scientific justification for the genocide during World War II. Owing to these historical
tragedies, some fears that a dominant focus on modern genetic models of mental illness could lead to exacerbated discrimination, and ultimately to history repeating itself.\textsuperscript{161}

Researchers must be aware of these historical misconducts, and must also consider the future possibilities and ethically challenging issues related to modern genetics research. In relation to the research questions and aims in this thesis, the ethical considerations related to the predictive value of genotyping are central. Given that, in the future, we may be able to predict psychological treatment responses, we also need to have a strategy for patients with a higher risk of non-response. The right to equal appropriate psychiatric health care, regardless of individual predictive factors, is fundamental.

5 CONCLUSION

This thesis presents the first findings of an association of aggregated genetic risk scores and CBT outcome in depression. Future studies will require larger and more homogenous samples to replicate the initial findings to untangle the overall genetic architecture involved in psychological treatment response.
6 ACKNOWLEDGEMENTS

Christian Rück. Where to start… Thank you. Thank you for being the most visionary, progressive, kindest, most scientifically decent and funniest main supervisor. Always having the patients and participants best in mind, and always, always having your PhD-student’s backs. I am so proud and thankful for being a part of your lab.

Cattis Lavebratt. Thank you for being my genetic co-supervisor, there have been a few significant times throughout this PhD-journey where your kindness, knowledge and support have been absolutely fundamental to me. You are one of the smartest, most hard working and good-hearted persons I have ever met.

Erik Hedman-Lagerlöf. As my one and only psychologist supervisor, I am forever thankful for your knowledge, generosity and contagious optimism, both professionally and as a friend. Always extremely funny, never judgmental.

Martin Schalling. The one and only Martin Schalling! Thank you for always having my back and generously share your long experience in the genetics field, and for inviting this rookie psychologist into your great lab at CMM.

Nils Lindefors. Thank you for letting me have the chance to build this research within your visionary and research-friendly organization, and thank you for always backing new research ideas in the most friendly of ways.

Eva Serlachius. Thank you Eva for being such an inspiration to younger (especially female) researchers. I have felt your support and I hope that we can continue the mentorship into future adventures, because now I need your wisdom and wit more than ever!

James Crowley. My pseudo-supervisor and GWAS go-to superperson. Thank you for always patiently answer all stupid questions in the kindest and most supportive of ways. It was a crucial moment in my PhD-education to have you on board and ever since that moment everything has been easy and fun (..sort of, kind of :).

Viktor Kaldo. I am forever thankful for you being the first person at Internetpsykiatrin (alongside Jan Bergström – thank you, too!) to reply to my email in 2008, saying that I was welcome into the experimental Internetpsykiatri lab at Huddinge Sjukhus. You manage to balance great scientific integrity with an always supporting approach and a great kindness.

1st and 2nd generation Rücklab. Completely unbiased, I thank the best research group on Earth. Fun, creative, supportive - all of you. Erik Andersson, Jesper Enander, Volen Ivanov – this
crew is my crew. So many fun and important moments we have shared, thank you! Julia Boberg, you fun and smart and witty research sister! I wish we will continue work and hang out for a long time to come (especially since you turn out to be such an awesome designer of Tables and figures!). Andrew Gentile, I am so glad you chose Huddinge Sjukhus over NYC so we can chat about research and American politics. Oscar Flygare, Lina Lundström, Diana Radu-Djurfeldt– thank you for always being generous, fun and supportive. And thank you rucklab hangarounds for always being great to be around - Josef Isung, Long-Long Chen, Adel Adu Hamdeh and Manuel Mattheisen.

Pat Sullivan. My other genetic/scientific go-to person/oracle – thank you for being such an inspiring and supportive researcher and person. Thank you for taking your time to help me out and for being so encouraging and making a junior psychiatric genetics researcher feel welcome and appreciated in the field. Almost too good to be true that you also have a great taste in movies!

Brjánn Ljótsson a.k.a mixed models guru. Apart from being one of the smartest persons alive, you are always immensely supportive, encouraging and fun. Thank you for being a great research colleague and friend.

Samir El Alaoui. Friend and go to person! Thank you for always being so supportive and immensely generous with your time and knowledge. You are one of a kind, both as a colleague and friend.

Kaldo lab. Kerstin Blom you blomma of a person! I have learned so many things from you it would be ridiculous to even start list them, but most of all you are smart, strong and so much fun, you will not get rid of me anytime soon. Berkeh Nasri – you are another inspiring research sister and friend, thank you for being you. Stay strong, sisters! Thank you rest of Kaldo lab for always being great fun and smart heads to be around; Susanna J, Martin K, Erik F, Simon M and all of the rest of you.

Monica Hellberg. All of the hundreds of papers published on participants from Internetpsykiatrienheten, owe a significant part of the credit to you and your tireless and thorough work. A true hero, that never takes any credit for it - thank you!

Nina Lind and Internetpsykiatrienheten. Thank every single one of you for keeping the data coming!
Annika Eriksson & Anna-Lee Jansen and the CMM lab. Thank you for making me feel so comfortable in your lab and for making it easier for a psychologist to wear the white lab coat! Thanks to all of the lovely people on CMM for making me feel welcome.

OCD team at G22. Thank you for the many OCD sessions and Christmas parties and for constant great fun and scientific support - David Mataix-Cols, Lorena Fernandez De la Cruz, Kayoko Isomura, Gustav Brander and the rest of the funny, smart and hard-working OCD squad!

Thank you friends for emotional support throughout these years. You know who you are!

Thank you mum Birgitta and dad Pehr, and to my in laws Roar and Anne-Ruth for always supporting me and for the incredible amount of babysitting this thesis required from all of you. It might be an obvious understatement to say but - I could not have done this without you. Thank you!!

My troop. Ingrid and Liv – nu är det över, nu ska vi ha kul igen och mamma ska sluta ”lata sig framför datorn”! Pål Sverre, husband – thank you for helping me do this every step of the way. “Let me count the ways”…

The 4266 participants who contributed directly to these studies. Thank you for trusting us and contributing to science in the most fundamental way. We owe everything to you.
7 REFERENCES


12. NICE. National Institute for Health and Clinical Excellence. Generalised anxiety disorder and panic disorder (with or without agoraphobia) in adults: management in primary, secondary and community care(Clinical guideline 113.)  


93. Major Depressive Disorder Working Group of the Psychiatric GC, Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM *et al*. A mega-analysis of genome-wide


