Worrying about the future: towards evidence-based prognosis in anxiety disorders.

Bokma, Wicher Alle

2021

document version
Publisher's PDF, also known as Version of record

WORRYING ABOUT THE FUTURE:

towards evidence-based prognosis in anxiety disorders

Wicher Alle Bokma

Amsterdam UMC
Universitair Medische Centra
department of psychiatry

WORRYING ABOUT THE FUTURE: TOWARDS EVIDENCE-BASED PROGNOSIS IN ANXIETY DISORDERS.

Wicher Alle Bokma

Amsterdam UMC
Universitair Medische Centra

Colofon

The infrastructure of the Netherlands Study of Depression and Anxiety (NESDA) study (www.nesda.nl) is funded through the Geestkracht program of the Netherlands Organization for Health Research and Development (ZonMw, grant number 10-000-1002) and is further supported by participant universities and mental health care organizations (VU University Medical Center [currently Amsterdam UMC location VUmc], GGZ ingest, Arkin, Leiden University Medical Center, Leiden University, GGZ Rivierduinen, University Medical Center Groningen, University of Groningen, Lentis, GGZ Friesland, GGZ Drenthe, Rob Giel Onderzoekscentrum).

The cardiac emergency department study was supported by a grant (project number 1103-056) from the Dutch organization Fonds Nuts Ohra.

Financial support for publication and distribution of this thesis was kindly provided by: Department of Psychiatry, Amsterdam University Medical Centers & VU University Amsterdam.

ISBN: 978-94-93270-17-6

Cover design: Tyler Hobbs - DECOMPILATION 5, reprinted with permission

Illustrations: Martin van den Berg

Layout & print: Guus Gijben | proefschrift-aio.nl

VRIJE UNIVERSITEIT

WORRYING ABOUT THE FUTURE: TOWARDS EVIDENCE-BASED PROGNOSIS IN ANXIETY DISORDERS.

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. C.M. van Praag,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Geneeskunde
op vrijdag 10 december 2021 om 13.45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

door

Wicher Alle Bokma

geboren te Amersfoort

promotoren:	prof.dr. A.J.L.M. van Balkom prof.dr. B.W.J.H. Penninx
copromotor:	dr. N.M. Batelaan
promotiecommissie:	prof.dr. A.T.F. Beekman prof.dr. J.W.R. Twisk dr. B.F.P. Broekman prof.dr. E.M. Sijbrandij prof.dr. P.H. Spinhoven prof.dr. K. Schruers

Table of contents

Chapter 1.

General introduction

Symptomatology

Anxiety disorders form a group of psychiatric disorders which are characterised by excessive psychological and physical anxiety responses to different subjectively threatening situations. In nature, the emotions of fear and anxiety are not pathological in many situations. Fear is what we experience when confronted with an acute threatening situation, like encountering a dangerous predator. Anxiety refers to a fear-like response in anticipation of a non-acute impending threatening situation. For example, anxiety can develop if we wander in an isolated, dangerous place in which predators or natural dangers might be looming around the corner. If we are presented with acute threats a fear response is an adaptive way of dealing with it as it prepares us for quick, decisive action.

Psychological anxiety responses include worrying, restlessness, concentration difficulties, irritability and sleep disturbances. Physical anxiety responses include increased respiration rate, muscle tension, fatigue, increased cardiac activity with palpitations, blushing, sweating, tingling sensations, headaches, abdominal or thoracic pain, dizziness, trembling, shaking, nausea and shortness of breath. In anxiety disorders, fear or anxiety responses are excessive, persisting and exist in association with non-threatening cues and are thus considered pathological.[1] Classification of different distinct anxiety disorders is based on the prevailing symptoms and on the provoking cues or situations. The most widely used classification manual is the Diagnostic and Statistical Manual, fifth edition (DSM-5).[1] The DSM-5 recognizes seven main distinct anxiety disorders: (1) Separation Anxiety Disorder, (2) Selective Mutism, (3) Specific Phobia, (4) Generalized Anxiety Disorder (GAD), (5) Panic Disorder (PD), (6) Social Anxiety Disorder (SAD, formerly known as social phobia), and (7) Agoraphobia. The current thesis will focus on the latter four anxiety disorders as these are the most prevalent and disabling anxiety disorders and as these four were included in the large dataset that will be used in this thesis for data analysis.

GAD is characterised by chronic anxiety and being overly concerned over everyday matters, in which catastrophic outcomes are dreaded. PD is characterised by recurring panic attacks and subsequent development of anticipatory anxiety for suffering more panic attacks. SAD is characterised by fear and anxiety in social interactions over public embarrassment, humiliation, and excessive concerns over being socially incompetent. All excessive anxiety responses are often followed by further non-adaptive

behavioural changes. A lot of patients will start to avoid situations or places in which their anxiety responses were previously triggered. Someone with PD might start to avoid places in which panic attacks were triggered. Someone with SAD might start avoiding social gatherings or situations in which they feel vulnerable to being judged. Persons with agoraphobia avoid public places or situations from which there are no escape options (*e.g.* public transportation, cinemas, theatres etc.) because they fear fainting, death or loss of control. Agoraphobia used to be a specifier for PD, but is considered a separate anxiety disorder since the introduction of the Diagnostic and Statistical Manual, fifth edition (DSM-5). See table 1 for the classification criteria for GAD, PD, SAD and Agoraphobia according to the DSM-5.[1]

Table 1. Criteria for classification of four anxiety disorders according to the Diagnostic and Statistical Manual, fifth edition.

	GAD	SAD	Agoraphobia	PD
Key features	Excessive anxiety and worry occurring more days than not about a number of events or activities. The individual finds it difficult to control the worry. Presence of three (or more) accompanying symptoms 1. Restlessness or feeling keyed up or on edge. 2. Being easily fatigued. 3. Difficulty concentrating or mind going blank. 4. Irritability. 5. Muscle tension. 6. Sleep disturbance	Marked fear or anxiety about one or more social situations in which the individual is exposed to possible scrutiny by others. Examples include social interactions (e.g., having a conversation, meeting unfamiliar people), being observed (e.g., eating or drinking), and performing in front of others (e.g., giving a speech).	A marked fear or anxiety about two (or more) of the following five situations: 1. Using public transportation 2. Being in open spaces 3. Being in enclosed spaces 4. Standing in line or being in a crowd 5. Being outside the home alone.	Recurrent unexpected panic attacks. A panic attack is an abrupt surge of intense fear or intense discomfort that reaches a peak within minutes, and during which time four (or more) of the following symptoms occur: 1. Palpitations, pounding heart, or accelerated heart rate. 2. Sweating. 3. Trembling or shaking. 4. Sensations of shortness of breath or smothering. 5. Feelings of choking. 6. Chest pain or discomfort. 7. Nausea or abdominal distress. 8. Feeling dizzy, unsteady, light-headed, or faint. 9. Chills or heat sensations. 10. Paresthesias 11. Derealization or depersonalization. 12. Fear of losing control or "going crazy." 13. Fear of dying.

Table continues

Table 1. Continued

	GAD	SAD	Agoraphobia	PD
Secondary features			The individual fears or avoids these situations because of thoughts that escape might be difficult or help might not be available in the event of developing panic-like symptoms or other incapacitating or embarrassing symptoms.	At least one of the attacks has been followed by one or both of the following: Persistent concern or worry about additional panic attacks or their consequences. A significant maladaptive change in behaviour related to the attacks
Specifier		Performance only: if the fear is restricted to speaking or performing in public.		
Duration	The fear, anxiety, or avoidance is persistent, typically lasting 6 months or more.			Panic attacks and subsequent worry or avoidance behaviours were present for at least 1 month or more
Out-of-proportion		The fear or anxiety is out of proportion to the actual threat posed by the social or agoraphobic situations and to the sociocultural context.		
Consistency		The social or agoraphobic situations almost always provoke fear or anxiety.		
Avoidance		The social or agoraphobic situations are avoided or endured with intense fear or anxiety, or require the presence of a companion.		
Distress/ impairment	The condition cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.			
Exclusion criteria	The disturbance is not attributable to the physiologic effects of a substance (e.g., a drug of abuse, a medication) or another medical condition (e.g., hyperthyroidism). If another medical condition is present, the fear, anxiety, or avoidance is clearly unrelated or is excessive. The fear, anxiety, or avoidance disturbance is not better explained by another mental disorder			

Adapted from: the Diagnostic and Statistical Manual, 5 edition.

Epidemiology

Excessive anxiety responses were already depicted in early historical texts on mental health. The earliest descriptions of excessive anxiety responses date back to Greek and Roman times.[2] In his 1638 book "The Anatomy of Melancholy" Robert Burton – under the pseudonym Democritus Junior – poetically described individuals who seemingly suffer from an anxiety disorder:

> *Many lamentable effects this fear causeth in men, as to be red, pale, tremble, sweat, it makes sudden cold and heat to come over all the body, palpitation of the heart, syncope (…). Many men are so amazed and astonished with fear, they know not where they are, what they*

> *say, what they do, and that which is worst, it tortures them many days before with continual affrights and suspicion. It hinders most honourable attempts, and makes their hearts ache, sad and heavy.* [3]

Nowadays, anxiety disorders are the most prevalent psychiatric disorders. 12-Month prevalence estimates for anxiety disorders (including specific phobia) in the general population range from 8.4-21.3%, whereas lifetime prevalence estimates range between 14.5-33.7%.[4-7] The age of onset tends to be during the formative years of adolescence and young adulthood: the median age of onset for SAD is 13 years, for agoraphobia is 20 years, for PD is 24 years and for GAD is 31 years.[8] Anxiety disorders are roughly twice as prevalent in females compared to males.[9] Anxiety disorders are more likely to develop in persons with high levels of psychological vulnerability: prevalence of anxiety disorders was increased around twofold in persons with premorbid high levels of neuroticism or anxiety sensitivity.[10,11] Patients with anxiety disorders suffer from substantial disability as a result of anxiety responses and accompanying avoidance behaviours. Globally, anxiety disorders are the seventh leading cause of years lived with disability; *i.e.* years of life lived in less than ideal health.[12] This makes the impact of anxiety disorders on global disability higher when compared with diabetes, migraine, asthma and ischemic heart disease. Critically, only a minority of patients with anxiety disorders seek professional help. In a large European general population study, only 20.6% of patients with an anxiety disorder reported ever receiving treatment.[13] Furthermore, patients with anxiety disorders are prone to high levels of comorbidity. Comorbidity with other psychiatric disorders is present in 85% of patients with GAD, 80% of patients with PD, 74% of patients with SAD and in 97% of patients with agoraphobia in the general population.[14] Comorbidity of two or more anxiety disorders is common. Correlations between PD with agoraphobia and SAD with agoraphobia are especially high, with tetrachoric correlations of 0.64 and 0.68 respectively, indicating that these disorders very often co-occur.[14] Among other psychiatric disorders, depressive disorders are the most frequent comorbid disorders. The correlations with anxiety disorders range between 0.43 and 0.62 for major depressive disorder, between 0.44-0.55 for dysthymia, between 0.43-0.49 for posttraumatic stress disorder (PTSD), between 0.38-0.51 for attention deficiency hyperactivity disorder (ADHD) and between 0.27-0.44 for any substance use disorder (SUD).[14] These correlations indicate that comorbidity with these disorders often occurs. The relative risk of having any chronic somatic disease such as hypertension, arthritis, asthma, ulcers, diabetes or cardiovascular diseases

is increased around two-fold for patients with anxiety disorders.[15] This is especially problematic as levels of disability are even further increased when a person with a chronic somatic disease also has a comorbid anxiety disorder.[16] Furthermore, economic costs associated with anxiety disorders are high. From scarcely available studies, it appears that in Europe yearly total additional costs associated with anxiety disorders range from €1,450 - €1,630 per patient.[17] A substantial portion of disability and economic costs associated with anxiety disorders is due to reduced work performance. The yearly work loss associated with anxiety disorders is estimated to be 17.6 days.[18] On the whole, anxiety disorders are highly prevalent disorders that are clearly associated with poor health outcomes, increased disability and high societal impact.

Pathophysiology

Genes

The aetiology of anxiety disorders is multicausal. First, there is strong evidence from twin studies that genetic factors underlie the development of anxiety disorders. The heritability estimates for GAD range between 32-49%, for SAD range between 39-56%, for PD equals 48%, and for Agoraphobia equals 67%.[19,20] This indicates that between one and two thirds of contributing factors in the aetiology of anxiety disorders are genetically defined. High levels of heritability exist in disorders in which genetic causes are associated with the development of this disorder at the population level. Many researchers devoted time and resources to unravel this genetic vulnerability by aiming to identify specific genes that contribute to the development of anxiety disorders. However, the candidate genes that were derived from this type of genetic research were only inconsistently associated with anxiety disorders and could not account for the high levels of heritability found in twin studies. This disparity between high heritability levels and low number of genes with a causal relationship to a disorder is termed the "missing heritability problem". Newer genetic research focuses on genome wide association studies (GWAS), as these methods have less risks of yielding inconsistent results. As of now, large GWAS studies into anxiety disorders showed only a small number of genetic aberrations, in the form of single nucleotide polymorphisms (SNPs).[21] These SNPs explain a mere 10% of the variance of the genetic vulnerability. In other words: 90% of underlying genetic susceptibility genes are still unknown. However, current GWAS efforts are still relatively underpowered when taking the low

hazard ratio of potential SNPs into account. Hopefully, future meta-analyses of GWAS findings will succeed in identifying additional susceptibility genes that are relevant to anxiety disorder pathophysiology and will thereby close the gap of missing heritability. Furthermore, future GWAS studies might expand beyond the classic case-control design into a dimensional approach to take heterogeneity of the anxiety disorder sample into account. This heterogeneity is underscored by the significant overlap in GWAS findings between anxiety disorders and the psychological trait neuroticism.[11] This highlights the possibility that a transdiagnostic approach into vulnerability factors for anxiety disorders instead of diagnostic entities such as anxiety disorder diagnoses could yield more meaningful genetic associations. The genetic contribution is in any case most likely polygenetic: not any single genetic factor determines all risk for development of anxiety disorders, but rather an interplay between many different genetic factors determines the risk.

Additionally, there are strong indicators that gene-environment (GxE) interactions are relevant in development of anxiety disorders.[22] In GxE interactions a genetic vulnerability for development of a disease will only lead to development of this disease when the individual encounters certain environmental stressors. In GxE interaction research, environmental stressors studied include stressful life events and childhood trauma. When these environmental stressors are encountered, a genetic vulnerability might lead to a higher tendency for anxiety responses. Furthermore, epigenetic processes likely contribute to the development of anxiety disorders. The term epigenetics is used to describe potentially heritable and functionally relevant modifications in gene expression without changes to the genetic code that are embedded within the DNA. These changes in gene expression can occur as a result of DNA-methylation. It is argued that certain environmental factors, such as childhood trauma or stressful life-events might evoke epigenetic changes.[23] However, as there are few candidate genes found that are relevant in the development of anxiety disorders, the relevance of GxE interaction and epigenetics is still largely unknown.[23]

Brain

The brain anatomy of patients with anxiety disorders shows differences in comparison with controls. For instance, patients with GAD have a larger grey matter volume in the amygdala.[20,24,25] Structural differences in amygdala, hippocampus, parahippocampal gyri, and brainstem nuclei are present in PD.[20,26] Brain anatomy findings in SAD were inconsistent.[27] The functions

of the associated brain regions are all related to emotional responses or processes. Besides structural brain anatomy differences, functional neuronal differences are present in anxiety disorders. These functional differences represent aberrant activation patterns in brain regions as a response to external stimuli. Usually, in laboratory settings, anxiety provoking tasks are used when studying functional neuronal differences. When confronted with anxiety provoking stimuli, anxiety disorder patients show greater amygdala activation and disruptions in amygdala-based intrinsic functional networks. This specific aberrant pattern of activation was named the 'fear circuit' due to its relevance to fear and anxiety responses.[24] Some other brain regions, such as the anterior cingulate cortex, the fusiform gyrus, the inferior frontal gyrus, the superior temporal gyrus, the globus pallidus and the insula also show different activation patterns in anxiety disorder patients.[24,25] Finally, some neuronal differences in anxiety disorder patients were shown when using novel neuroimaging techniques, such as PET, SPECT and metabolic MRI.[25] These findings are still too inconsistent and should be considered preliminary.

Understanding neuroimaging differences in anxiety disorders are important as they could potentially provide a rationale for certain types of treatments. For instance, if a certain drug type would be able to mitigate the aberrant activation in the fear circuit in anxiety disorders that drug could prove effective in treatment of anxiety disorders. Also, neuroimaging findings might be indicative of certain subtypes or differences in clinical course. However, currently these neuroimaging findings cannot yet be translated into daily clinical practice.[28]

Neurochemistry

The cornerstone of anxiety disorders is the heightened emotional response after a subjectively threatening stimulus. Both the autonomic nervous system (ANS) and the hypothalamic-pituitary-adrenal axis (HPA-axis) are activated in this response.[29] The response of the ANS consists of increased sympathetic activation and decreased parasympathetic activation. Both result in changes in various bodily functions: control of respiration, cardiac regulation (the cardiac control centre), vasomotor activity (the vasomotor centre), and certain reflex actions such as coughing, sneezing, swallowing and vomiting. Many of these autonomic effects are indeed present when patients experience anxiety responses. Furthermore, it seems that aberrant breathing patterns and reduced heart rate variability are present in PD patients even if they are not currently experiencing panic symptoms.[26,30]

These findings likely reflect ongoing changes in ANS activity in PD patients. These ongoing autonomic disruptions lead to lingering somatic anxiety phenomena and can reduce the threshold for triggering a full-blown panic attack. In this study the effects were adjusted for use of psychotropic medication that are known to influence the heart rate variability. However, there still is debate as of whether these changes are present as the largest single study into autonomic dysregulation in anxiety disorders did not show differences in resting state autonomic function between anxiety disorders and controls.[31]

It is a longstanding theory that the HPA-axis is involved in anxiety disorders, as the main physiologic functions of the HPA-axis revolve around stress responses.[26] Indeed, patients with anxiety disorders show overactivation of the HPA-axis.[32,33] Higher HPA axis activation leads to increased levels of cortisol, the stress hormone, which triggers several somatic responses that can be anxiety provoking. This heightened stress response might lead to non-threatening situations being perceived as critically threatening by patients with anxiety disorders. Probably, heightened HPA-axis activity is the pathway between childhood trauma and development of anxiety disorders, as childhood trauma is a strong risk factor for both adult presence of psychiatric disorders and for increased HPA-axis activity at adulthood.[34] However, some studies show reduced levels of cortisol in presence of anxiety symptoms.[35] Therefore, the direction of causality is unsure and the results from this field of research are still inconclusive.

Biochemistry

Different neurotransmitters are associated with anxiety disorders. Plasma levels of 5-hydroxytryptamine (serotonin, 5-HT) and gamma aminobutyric acid (GABA) activity were linked to GAD.[25] These signalling hormones play a vital role in communication within and between different brain regions. Many anti-anxiety drugs target these neurotransmitter pathways. Besides altered plasma levels, PET studies are suggestive of changes in 5-HT and GABA neuronal circuitry involved in anxiety processing in anxiety disorders.[25] Other research into biological markers for anxiety disorders shows inconsistent findings. For instance, some studies showed that anxiety disorders are linked to higher CRP-levels,[36–38] while others found no association.[39] Likewise, anxiety symptoms were also linked to both higher[33] and lower cortisol,[35] as well as higher [35,37,40] and lower interleukin-6 (IL-6) measurements.[36] Furthermore, inconsistent findings were reported for associations with metabolic syndrome markers,[41–45] tumour necrosis factor-α (TNF-α)

levels,[37,40] and BDNF levels.[46,47] In spite of decades of rigorous scientific work there are currently no clinical tests based on biological differences available for psychiatric disorders.[28]

Clinical course

Once anxiety disorders have developed, clinical course is heterogenous: while some patients fully recover, others develop chronic symptoms. A naturalistic 12-year study in a clinical sample showed that recovery rates during the 12-year period varied across diagnoses: 82% of PD patients recovered, while recovery occurred in only 58% of GAD patients, in 48% of PD patients who had agoraphobia, and 37% of SAD patients.[48]

When left untreated, anxiety disorders might remit, but more likely symptoms will remain persistent. A study assessing one-year follow-up on anxiety disorder severity showed that symptomatology hardly changed in untreated samples.[49] Unfortunately, many patients with anxiety disorders do not seek professional care. Around 80-90% of patients with anxiety disorders do not access professional care or delay accessing it.[13,50] A delay of 15 years for seeking professional care is not at all out of the ordinary. Even in presence of evidence-based treatments, many patients do not fully remit. A 12-year naturalistic follow-up study in a clinical sample showed that probabilities of experiencing eight consecutive weeks of having no or mild anxiety symptoms were modest.[48] Although it is clear that a substantial number of anxiety disorder patients do not benefit from different treatments, no clear definition for treatment resistance in anxiety disorders exists.

Several individual risk factors for chronicity or suboptimal treatment results in anxiety disorders are known: higher baseline severity of anxiety symptoms, presence of comorbidity, higher levels of disability, longer previous duration of anxiety symptoms, younger age of onset, longer duration of untreated anxiety symptoms, and presence of childhood trauma.[48,51–55] Also, sociodemographic and lifestyle factors such as lower education years, higher age, having no partner, having low levels of social support, smoking and nicotine dependency, having financial problems, and being unemployed or having a low income were associated with poor outcomes in anxiety disorder patients.[48,49,55–58] Additionally, psychological traits such as high neuroticism, high anxiety sensitivity, high levels of worrying, low extraversion, and low levels of mastery are related to poor outcomes in anxiety disorders.[51,55,59] Although some individual risk factors are identified, due to the complex

interplay between these factors much is still unknown about how clinical course in anxiety disorders is defined. As a result, these risk factors cannot yet be implemented in clinical care or in clinical decision making.

Treatment

Psychotherapy like cognitive behavioural therapy (CBT) is shown to be effective in treating anxiety disorders, with large effect sizes: Hedges *g* was 0.80 for CBT versus control conditions in GAD, 0.81 for CBT versus control conditions in PD, and 0.88 for CBT versus control conditions in SAD.[62] In this meta-analysis of 31 studies, control conditions consisted of waiting lists (n=24), care as usual (n=4) and pill placebos (n=3). Generally, treatment results were largest when comparisons were made against waiting lists and were smaller, but still significant, when comparisons were made against care as usual or pill placebo. This comes as no surprise because a waiting list control condition only controls for natural course, pill placebo for natural course and nonspecific treatment effects and care as usual for natural course as well as nonspecific and specific effects treatment effects. Furthermore, pharmacotherapy is effective in treatments of anxiety disorders. Different classes of medication are used in anxiety disorders, *e.g.* selective serotonin reuptake inhibitor (SSRIs), serotonin-noradrenaline reuptake inhibitors (SNRIs), benzodiazepines (BZDs), tricyclic antidepressants (TCAs), tetracyclic antidepressant, monoamine oxidase inhibitors (MAOIs), anticonvulsants and antipsychotics. Due to the likely involvement of serotonergic pathways in anxiety disorders, a preference for serotonergic agents exists. These include SSRI's, as well as SNRIs, the TCAs clomipramine and imipramine and the MAOI phenelzine. All of these medication classes show a significant effect in anxiety disorders, with Cohen's *d* pre- and posttreatment ranging from 1.83 for TCAs to 2.25 for SNRIs.[63] These effect sizes indicate a very large beneficial effect of initiating a pharmacologic treatment regimen. In comparison with pill placebo, the effects of pharmacotherapy in anxiety disorders remain substantial, with effect sizes around 0.60 for SSRIs and 0.50 for SNRIs.[64] Careful consideration is needed in treatment selection as pharmacotherapeutics have side-effects. SSRIs are considered first-line treatments in all anxiety disorders as these are usually tolerated best.

Prognosis

The cornerstones of clinical practice in medicine consist of diagnosis, prognosis and treatment. The relevance of prognosis for clinicians was first recognized by Hippocrates, who believed that a valid prognosis should follow

from the physician's assessment of an individual patient, using knowledge on pathophysiology and on factors that define the clinical course.[65,66] However, the prognoses provided by physicians are not empirically validated. For instance, experienced radiology oncologists who specialized in lung cancer performed poorly when predicting two-year deaths, dysphagia and dyspnoea in lung cancer patients.[67] Emergency Department physicians were reasonably able to predict new occurrences of asthma exacerbations in asthmatic patients.[68] Psychiatrists were reasonably able to predict recurring suicidal behaviour in the next 6-months in patients they assessed for suicidal behaviour.[69] They were, however, unable to predict 6-month incidence of suicide in the same patients. These examples illustrate the lack of accuracy in clinician opinion prognoses in medicine. Instead of clinician opinion prognoses, prediction models can be used to provide individual prognoses. In prediction models, the prognosis is based on statistical associations between individual patients and disease characteristics with outcomes.[70] There is ample evidence that statistical prediction methods can improve poor predictive properties of clinician opinion prognoses.[66,69,71]

Precision psychiatry

In spite of the historical awareness of the importance of prognosis, the science around prognosis in medicine is still lacking.[72] Prognosis in psychiatry is traditionally based on classifications, for instance as described in the DSM-5, and as provided in Table 1 with regard to anxiety disorders (see above). In these classification systems, psychiatric syndromes are delineated on the basis of clusters of symptoms that often co-occur. They do not classify on the basis of pathophysiologic characteristics. They are not meant to demarcate an underlying 'disease' but rather describe symptoms in a standardized way. Although the use of classification systems like the DSM-5 and its predecessors was vital for the development of the field of psychiatry, their classifications proved insufficient to base prognoses on. For instance, DSM-IV anxiety disorder classifications yield less precise course predictions in comparison to clinical characteristics such as severity of symptoms, duration of symptoms and level of disability.[57] Improving course prediction would be an important step towards personalized medicine, which aims to individually tailor diagnosis and prognosis based on individual disease factors to derive a personalized treatment plan. Operationalizations of personalized medicine in psychiatry are often touted as "precision psychiatry". Providing reliable prognoses is a vital aim for precision psychiatry. Prediction models seem promising for improving evidence-based prognosis and thereby further evolve the field of precision psychiatry. Prediction models aim to combine

different types of data to predict a future outcome.[70] Different statistical methods can underlie such models, for instance machine learning algorithms show a lot of promise in deriving prediction models in psychiatry.[73]

As is clear from this introduction, due to the high levels of chronicity and disability in anxiety disorders, adequate identification of anxiety disorder patients with higher risk profiles is much needed in clinical care. In spite of substantial knowledge on pathophysiology of anxiety disorders and risk factors for clinical course in anxiety disorders, it currently remains impossible to adequately predict the disease course in individual patients. In other words, it is time for the development of evidence-based prognosis in psychiatry.

Aims of this thesis

This thesis aims to contribute to the development of evidence-based prognosis in two ways. The first aim is to increase knowledge on factors that impact the clinical course in anxiety disorders. The second aim is to use existing knowledge on pathophysiology and risk factors for anxiety disorders to predict clinical course in anxiety disorders over time.

Sample and study design

For a number of chapters in this thesis (chapters 3, 5, 6 and 7), participants were recruited from the Netherlands Study of Depression and Anxiety (NESDA), a multi-centre naturalistic longitudinal cohort study among adult respondents (aged 18-65) from different regions in the Netherlands. Respondents were recruited from the community, primary care and specialized mental health care settings and the sampling was stratified to be representative of the various developmental stages of depression and anxiety. At baseline, 2,981 subjects were included. The main aim of NESDA is to gain insight into the long-term course and consequences of anxiety and depressive disorders. Baseline assessments were conducted at the three participating sites between 2004 and 2007. Follow-up measurements were performed at one-year, two-year, four-year, six-year, and nine-years after baseline. NESDA represents ongoing research as currently thirteen-year follow-up measurements are being performed (2019-2022). Baseline measurements included sociodemographics, clinical characteristics, psychological assessments, biological assessments and structured psychiatric interviews assessing DSM-IV diagnoses. Comorbid psychiatric

disorders were permitted with the exception of psychotic disorders, bipolar disorders, PTSD, obsessive-compulsive disorder, or severe SUDs, as reported by participants or their mental health care practitioner. Participants were excluded if they showed insufficient proficiency in the Dutch language. The Ethical Committee of participating universities approved of the study protocol and all participants provided their written informed consent.

Contents of this thesis

The first part of this thesis explores the first main aim to increase knowledge on factors that impact the clinical course in anxiety disorders. **Chapter 2** focusses on diagnosis and early detection of anxiety disorders in a general hospital sample. In this chapter, a screening programme for anxiety disorders was tested in the cardiac emergency department. We hypothesized that physical and psychological anxiety responses are present in a large number of non-cardiac chest pain patients and that these anxiety responses are indicative of underlying anxiety disorders, thereby providing an opportunity for early recognition and referral for psychiatric treatment. **Chapter 3** is a cross-sectional study into the effects of chronic somatic diseases on disability and work-loss in anxiety disorders and depressive disorders. In this chapter, the hypothesis that presence of comorbidity between anxiety disorders and chronic somatic diseases negatively impacts outcomes in anxiety disorders is tested. **Chapter 4** describes a narrative systematic review into different aspects of treatment resistance in anxiety disorders. In this way, factors that are deemed relevant in the development of chronicity in anxiety disorders are gathered. In this chapter, a proposal for a consensus definition for treatment resistance in anxiety disorders was formulated.

The second part of this thesis focuses on the second aim to predict clinical course in anxiety disorders over time. In this part of the thesis, a number of prediction models are developed in order to assess the predictive properties of combinations of different pathophysiologic factors and risk factors. These prediction models are presented in accordance to a methodological hierarchy: the first prediction model assessed is one that is based on clinician opinion; the second prediction model assessed is based on the results from a systematic review, and the third prediction model is a data-driven machine learning approach. **Chapter 5** presents a clinical staging model that consists of different stages that can be ordered as ordinal categories. In this chapter, a clinical staging model from a well-known Australian research group was adapted for use in anxiety disorders and the predictive properties were tested in a large sample consisting of at-risk controls and anxiety disorder patients

over a six-year timespan. **Chapter 6** presents a dimensional measurement instrument that can be ordered as interval categories. In this chapter, the results from our systematic review into definitions for treatment resistance in anxiety disorders (chapter 4) were incorporated into a measurement tool that assesses the degree of treatment resistance in anxiety disorders. The predictive properties for this measurement instrument were assessed in anxiety disorder patients who received treatments over a two-year period. **Chapter 7** presents a dichotomous prediction model. This is a data-driven approach in which machine learning methods were used to derive a prediction model for two-year outcomes in anxiety disorder based on various baseline measurements. The model is based on random forests classifiers using a wide array of baseline predictors. The predictive properties of these predictions were assessed over a two-year follow-up period.

Overall, this thesis aims to contribute to the development of evidence-based prognosis. More accurate course predictions in anxiety disorders could have significant implications for clinical care. It could lead to personalized risk assessments. Hopefully, improved course predictions can be used in clinical decision making: providing the right intervention at the right moment for the right patient.

References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.; 2013. doi:10.1176/appi.books.9780890425596.744053
2. Crocq MA. A history of anxiety: From Hippocrates to DSM. *Dialogues Clin Neurosci*. 2015;17(3):319-325. doi:10.31887/dcns.2015.17.3/macrocq
3. Democritus Junior. *The Anatomy of Melancholy*. The Gutenberg Project; 1638. https://www.gutenberg.org/files/10800/10800-h/10800-h.htm.
4. Bourdon KH, Rae DS, Locke BZ, Narrow WE, Regier DA. Estimating the prevalence of mental disorders in U.S. adults from the Epidemiologic Catchment Area Survey. *Public Health Rep*. 1992;107(6):663-668.
5. Alonso J, Angermeyer MC, Bernert S, et al. Prevalence of mental disorders in Europe: Results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand*. 2004;109(Suppl. 420):21-27. doi:10.1111/j.1600-0047.2004.00327.x
6. Wittchen HU, Jacobi F. Size and burden of mental disorders in Europe - A critical review and appraisal of 27 studies. *Eur Neuropsychopharmacol*. 2005;15(4):357-376. doi:10.1016/j.euroneuro.2005.04.012
7. Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen H-U. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *Int J Methods Psychiatr Res*. 2012;21(3):169-184. doi:10.1002/mpr
8. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):593-602. doi:10.1001/archpsyc.62.6.593
9. Bandelow B, Michaelis S. Epidemiology of anxiety disorders in the 21st century. *Dialogues Clin Neurosci*. 2015;17(327):335. doi:10.1016/j.siny.2015.10.004
10. Schmidt NB, Zvolensky MJ, Maner JK. Anxiety sensitivity: Prospective prediction of panic attacks and Axis I pathology. *J Psychiatr Res*. 2006;40(8):691-699. doi:10.1016/j.jpsychires.2006.07.009
11. Ormel J, Jeronimus BF, Kotov R, et al. Neuroticism and Common Mental Disorders: Meaning and Utility of a Complex Relationship. *Clin Psychol Rev*. 2013;33(5):686-697.
12. Vos T, Flaxman AD, Naghavi M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012;380(9859):2163-2196. doi:10.1016/S0140-6736(12)61729-2

13. Alonso J, Lépine J. Overview of key data from the European Study of the Epidemiology of Mental Disorders (ESEMeD). *J Clin Psychiatry*. 2007;68(suppl 2):3-9.
14. Kessler RC, Chiu WT, Demler O, Walters EE, Merikangas KR. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(June 2005):617-627. doi:10.1001/archpsyc.62.6.617
15. Roy-Byrne PP, Davidson KW, Kessler RC, et al. Anxiety disorders and comorbid medical illness. *Gen Hosp Psychiatry*. 2008;30(3):208-225. doi:10.1016/j.genhosppsych.2007.12.006
16. Sareen J, Cox BJ, Clara I, Asmundson GJG. The relationship between anxiety disorders and physical disorders in the U.S. National Comorbidity Survey. *Depress Anxiety*. 2005;21(4):193-202. doi:10.1002/da.20072
17. Andlin-Sobocki P, Wittchen HU. Cost of anxiety disorders in Europe. *Eur J Neurol*. 2005;12(SUPPL. 1):39-44. doi:10.1111/j.1468-1331.2005.01196.x
18. Buist-Bouwman MA, De Graaf R, Vollebergh WAM, Ormel J. Comorbidity of physical and mental disorders and the effect on work-loss days. *Acta Psychiatr Scand*. 2005;111(6):436-443. doi:10.1111/j.1600-0447.2005.00513.x
19. Mather L, Blom V, Bergström G, Svedberg P. An Underlying Common Factor, Influenced by Genetics and Unique Environment, Explains the Covariation between Major Depressive Disorder, Generalized Anxiety Disorder, and Burnout: A Swedish Twin Study. *Twin Res Hum Genet*. 2016;19(6):619-627. doi:10.1017/thg.2016.73
20. Bandelow B, Baldwin D, Abelli M, et al. Biological markers for anxiety disorders, OCD and PTSD – a consensus statement. Part I: Neuroimaging and genetics. *World J Biol Psychiatry*. 2016;17(5):321-365. doi:10.1080/15622975.2016.1181783
21. Otowa T, Hek K, Lee M, et al. Meta-analysis of genome-wide association studies of anxiety disorders. *Mol Psychiatry*. 2016;21(10):1391-1399. doi:10.1038/mp.2015.197
22. Musci RJ, Augustinavicius JL, Volk H. Gene-Environment Interactions in Psychiatry: Recent Evidence and Clinical Implications. *Curr Psychiatry Rep*. 2019;21(9). doi:10.1007/s11920-019-1065-5
23. Kim Y-K, ed. *Anxiety Disorders: Rethinking and Understanding Recent Discoveries.* Vol 1191. Singapore: Springer Nature; 2020. doi:10.1007/978-981-32-9705-0_28
24. Etkin A, Wager TD. Functional neuroimaging of anxiety: A meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *Am J Psychiatry*. 2007;164(10):1476-1488. doi:10.1176/appi.ajp.2007.07030504

25. Maron E, Nutt D. Biological markers of generalized anxiety disorder. *Dialogues Clin Neurosci*. 2017;19:147-158.
26. Cosci F, Mansueto G. Biological and clinical markers in panic disorder. *Psychiatry Investig*. 2019;16(1):27-36. doi:10.30773/pi.2018.07.26
27. Brühl AB, Delsignore A, Komossa K, Weidt S. Neuroimaging in social anxiety disorder-A meta-analytic review resulting in a new neurofunctional model. *Neurosci Biobehav Rev*. 2014;47:260-280. doi:10.1016/j.neubiorev.2014.08.003
28. Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry*. 2012;17(12):1174-1179. doi:10.1038/mp.2012.105
29. Rotenberg S, McGrath JJ. Inter-relation between autonomic and HPA axis activity in children and adolescents. *Biol Psychol*. 2016;117:16-25. doi:10.1016/j.biopsycho.2016.01.015
30. Chalmers JA, Quintana DS, Abbott MJ-A, Kemp AH. Anxiety Disorders are Associated with Reduced Heart Rate Variability: A Meta-Analysis. *Front psychiatry*. 2014;5(July):80. doi:10.3389/fpsyt.2014.00080
31. Licht CM, de Geus EJ, van Dyck R, Penninx BW. Association between anxiety disorders and heart rate variability in the Netherlands study of depression and anxiety (NESDA). *Psychosom Med*. 2009;71(5):508-518. doi:10.1097/PSY.0b013e3181a292a6
32. Faravelli C. Childhood stressful events, HPA axis and anxiety disorders. *World J Psychiatry*. 2012;2(1):13. doi:10.5498/wjp.v2.i1.13
33. Zoccola PM, Dickerson SS, Yim IS. Trait and state perseverative cognition and the cortisol awakening response. *Psychoneuroendocrinology*. 2011;36(4):592-595. doi:10.1016/j.psyneuen.2010.10.004
34. Tyrka AR, Wier L, Price LH, et al. Childhood Parental Loss and Adult Hypothalamic-Pituitary-Adrenal Function. *Biol Psychiatry*. 2008;63(12):1147-1154. doi:10.1016/j.biopsych.2008.01.011
35. O 'Donovan A, Hughes BM, Slavich GM, et al. Clinical anxiety, cortisol and interleukin-6: Evidence for specificity in emotion–biology relationships. *Brain Behav Immun*. 2010;24:1074-1077. doi:10.1016/j.bbi.2010.03.003
36. Vogelzangs N, Beekman ATF, De Jonge P, Penninx BWJH. Anxiety disorders and inflammation in a large adult cohort. *Transl Psychiatry*. 2013;3. doi:10.1038/tp.2013.27
37. Pitsavos C, Panagiotakos DB, Papageorgiou C, Tsetsekou E, Soldatos C, Stefanadis C. Anxiety in relation to inflammation and coagulation markers, among healthy adults: The ATTICA Study. *Atherosclerosis*. 2006;185(2):320-326. doi:10.1016/j.atherosclerosis.2005.06.001

38. Copeland WE, Shanahan L, Worthman C, Angold A, Costello EJ. Generalized anxiety and C-reactive protein levels: A prospective, longitudinal analysis. *Psychol Med*. 2012;42(12):2641-2650. doi:10.1017/S0033291712000554
39. O'Donovan A, Hughes BM, Slavich GM, et al. Clinical anxiety, cortisol and interleukin-6: Evidence for specificity in emotion-biology relationships. *Brain Behav Immun*. 2010;24(7):1074-1077. doi:10.1016/j.bbi.2010.03.003
40. Hoge EAA, Brandstetter K, Moshier S, Pollack MHH, Wong KKK, Simon NMM. Broad spectrum of cytokine abnormalities in panic disorder and posttraumatic stress disorder. *Depress Anxiety*. 2009;26:447-455. doi:10.1002/da.20564
41. Kahl KG, Schweiger U, Correll C, et al. Depression, anxiety disorders, and metabolic syndrome in a population at risk for type 2 diabetes mellitus. *Brain Behav*. 2015;5(3):e00306. doi:10.1002/brb3.306
42. Carroll D, Phillips AC, Thomas GN, Gale CR, Deary I, Batty GD. Generalized Anxiety Disorder Is Associated with Metabolic Syndrome in the Vietnam Experience Study. *Biol Psychiatry*. 2009;66(1):91-93. doi:10.1016/j.biopsych.2009.02.020
43. Perez-Cornago A, Ramírez MJ, Zulet MÁ, Martinez JA. Effect of dietary restriction on peripheral monoamines and anxiety symptoms in obese subjects with metabolic syndrome. *Psychoneuroendocrinology*. 2014;47:98-106. doi:10.1016/j.psyneuen.2014.05.003
44. Skilton MR, Moulin P, Terra J-L, Bonnet F. Associations Between Anxiety, Depression, and the Metabolic Syndrome. *Biol Psychiatry*. 2007;62(11):1251-1257.
45. Butnoriene J, Bunevicius A, Norkus A, Bunevicius R. Depression but not anxiety is associated with metabolic syndrome in primary care based community sample. *Psychoneuroendocrinology*. 2014;40(1):269-276. doi:10.1016/j.psyneuen.2013.11.002
46. Molendijk ML, Bus BA, Spinhoven P, et al. Gender specific associations of serum levels of brain-derived neurotrophic factor in anxiety. *World J Biol Psychiatry*. 2012;13(7):535-543. doi:10.3109/15622975.2011.587892
47. Suliman S, Hemmings SM, Seedat S. Brain-Derived Neurotrophic Factor (BDNF) protein levels in anxiety disorders: Systematic review and meta-regression analysis. *Front Integr Neurosci*. 2013;7(JUL):1-11. doi:10.3389/fnint.2013.00055
48. Bruce SE, Yonkers KA, Otto MW, et al. Influence of Psychiatric Comorbidity on Recovery and Recurrence in Generalized Anxiety Disorder, Social Phobia, and Panic Disorder: A 12-Year Prospective Study. *Am J Psychiatry*. 2005;162(6):1179-1187. doi:10.1176/appi.ajp.162.6.1179
49. van Beljouw IM, Verhaak PF, Cuijpers P, van Marwijk HW, Penninx BW. The course of untreated anxiety and depression, and determinants of poor one-year outcome: a one-year cohort study. *BMC Psychiatry*. 2010;10:86. doi:10.1186/1471-244X-10-86

50. Johnson EM, Coles ME. Failure and delay in treatment-seeking across anxiety disorders. *Community Ment Health J*. 2013;49(6):668-674.
51. Spinhoven P, Batelaan NM, Rhebergen D, van Balkom AL, Schoevers R, Penninx BW. Prediction of 6-yr symptom course trajectories of anxiety disorders by diagnostic, clinical and psychological variables. *J Anxiety Disord*. 2016;44:92-101. doi:10.1016/j.janxdis.2016.10.011
52. Hendriks SM, Spijker J, Licht CMM, Beekman ATF, Penninx BWJH. Two-year course of anxiety disorders: different across disorders or dimensions? *Acta Psychiatr Scand*. 2013;128(3):212-221. doi:10.1111/acps.12024
53. Beesdo-Baum K, Knappe S, Fehm L, et al. The natural course of social anxiety disorder among adolescents and young adults. *Acta Psychiatr Scand*. 2012;126(6):411-425. doi:10.1111/j.1600-0447.2012.01886.x
54. Baldwin DS, Tiwari N. The pharmacologic treatment of patients with generalized anxiety disorder: where are we now and where are we going? *CNS Spectr*. 2009;14(2 Suppl 3):5-12.
55. Asselmann E, Beesdo-Baum K. Predictors of the Course of Anxiety Disorders in Adolescents and Young Adults. *Curr Psychiatry Rep*. 2015;17(2):1-8. doi:10.1007/s11920-014-0543-z
56. Catarino A, Bateup S, Tablan V, et al. Demographic and clinical predictors of response to internet-enabled cognitive–behavioural therapy for depression and anxiety. *BJPsych Open*. 2018;4(5):411-418. doi:10.1192/bjo.2018.57
57. Batelaan NM, Rhebergen D, Spinhoven P, van Balkom AJ, Penninx BWJH. Two-Year Course Trajectories of Anxiety Disorders: Do DSM Classifications Matter? *J Clin Psychiatry*. 2014;75(09):985-993. doi:10.4088/JCP.13m08837
58. Nay W, Brown R, Roberson-Nay R. Longitudinal course of panic disorder with and without agoraphobia using the national epidemiologic survey on alcohol and related conditions (NESARC). *Psychiatry Res*. 2013;208(1):54-61. doi:10.1016/j.psychres.2013.03.006
59. Scholten WD, Batelaan NM, van Balkom AJ, Penninx BW, Smit JH, Van Oppen P. Recurrence of anxiety disorders and its predictors. *J Affect Disord*. 2013;147(1-3):180-185. doi:10.1016/j.jad.2012.10.031
60. Schiefelbein VL, Susman EJ. Cortisol levels and longitudinal cortisol change as predictors of anxiety in adolescents. *J Early Adolesc*. 2006;26(4):397-413. doi:10.1177/0272431606291943
61. Kobayashi K, Shimizu E, Hashimoto K, et al. Serum brain-derived neurotrophic factor (BDNF) levels in patients with panic disorder: As a biological predictor of response to group cognitive behavioral therapy. *Prog Neuro-Psychopharmacology Biol Psychiatry*. 2005;29(5):658-663. doi:10.1016/j.pnpbp.2005.04.010

62. Cuijpers P, Cristea IA, Karyotaki E, Reijnders M, Huibers MJH. How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. *World Psychiatry*. 2016;15(3):245-258. doi:10.1002/wps.20346
63. Bandelow B, Reitt M, Röver C, Michaelis S, Görlich Y, Wedekind D. Efficacy of treatments for anxiety disorders: A meta-analysis. *Int Clin Psychopharmacol*. 2015;30(4):183-192. doi:10.1097/YIC.0000000000000078
64. Jakubovski E, Johnson JA, Nasir M, Müller-Vahl K, Bloch MH. Systematic review and meta-analysis : Dose – response curve of SSRIs and SNRIs in anxiety disorders. *Depress Anxiety*. 2019;36(3):198-212. doi:10.1002/da.22854
65. Hippocrates. *The Book of Prognostics*. 400 B.C. Gloucester, England: Dodo Press; 2009.
66. Thompson DD, Murray GD, Sudlow CLM, Dennis M, Whiteley WN. Comparison of statistical and clinical predictions of functional outcome after ischemic stroke. *PLoS One*. 2014;9(10):1-7. doi:10.1371/journal.pone.0110189
67. Oberije C, Nalbantov G, Dekker A, et al. A prospective study comparing the predictions of doctors versus models for treatment outcome of lung cancer patients: a step towards individualized care and shared decision making. *Radiother Oncol*. 2014;112(1):37-43. doi:10.1016/j.radonc.2014.04.012.A
68. Farion KJ, Wilk S, Michalowski W, O'Sullivan D, Sayyad-Shirabad J. Comparing predictions made by a prediction model, clinical score, and physicians. *Appl Clin Inform*. 2013;4(3):376-391. doi:10.4338/ACI-2013-04-RA-0029
69. Randall JR, Sareen J, Chateau D, Bolton JM. Predicting Future Suicide: Clinician Opinion versus a Standardized Assessment Tool. *Suicide Life-Threatening Behav*. 2019;49(4):941-951. doi:10.1111/sltb.12481
70. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. (Gail M, Krickeberg K, Samet J, Tsiatis A, Wong W, eds.). New York, NY, US: Springer; 2009. doi:10.1007/978-1-4757-3294-8_4
71. Meehl PE. *Clinical versus Statistical Prediction*. Minneapolis: University of Minnesota; 1954.
72. Croft P, Altman DG, Deeks JJ, et al. The science of clinical practice: Disease diagnosis or patient prognosis? Evidence about "what is likely to happen" should shape clinical practice. *BMC Med*. 2015;13(1):1-8. doi:10.1186/s12916-014-0265-4
73. Iniesta R, Stahl D, McGuffin P. Machine learning, statistical learning and the future of biological research in psychiatry. *Psychol Med*. 2016;46(May):2455-2465. doi:10.1017/S0033291716001367

Chapter 2.

Feasibility and outcome of the implementation of a screening program for panic disorder in noncardiac chest pain patients in cardiac emergency department routine care

General Hospital Psychiatry (2015)

Wicher A. Bokma
Neeltje M. Batelaan
Aernout M. Beek
Annette D. Boenink
Jan H. Smit
Anton J.L.M. van Balkom

Abstract

Objective: This study assesses the feasibility and outcome of the implementation of a screening program for classifying Panic Disorder (PD) in patients presenting with non-cardiac chest pain (NCCP), when integrated in routine Cardiac Emergency Department (CED) care.

Methods: Barrier analyses were made during the pilot phase and implementation period. NCCP-patients aged 18-70 years presenting at the CED (n=252) were eligible for screening with the Hospital Anxiety and Depression Scale (HADS). Those scoring above cut-off on the HADS were referred to the Psychiatric Department and received the Composite International Diagnostic Interview.

Results: Screening was initiated in 60 patients (23.8%), of whom nine refused participation. Staff- adherence remained low despite implementing several improvements in the screening procedure. In total, 39 patients completed the program, 8 were diagnosed with a psychiatric disorder, including two patients with PD.

Conclusion: Feasibility of implementation of this screening program for PD in NCCP-patients in routine CED care was limited, because offering screening frequently conflicted with providing acute care, and because patients showed relatively high refusal rates. Contrasting our assumption, various other psychiatric disorders besides PD were classified.

Introduction

In 50-63% of patients with acute chest pain presenting at the Cardiac Emergency Department (CED), no cardiac cause is found for the complaints and 'non-cardiac chest pain' (NCCP) is diagnosed.[1,2] Panic Disorder (PD) is highly prevalent (12-41%) among NCCP-patients.[3,4] Symptoms of a panic attack may occur sudden and may mimic those of a heart attack.[5] In many PD patients presenting with NCCP, the diagnosis of PD is overlooked [1,6] and is left untreated.[7-9] When PD is recognized in NCCP-patients effective treatment regimes exist.[7,9] This study examines the implementation process, patient and staff-adherence, and outcome of a screening program aimed at integrating psychiatric screening in *routine* CED care of NCCP by identifying persons with PD. This is the first study to evaluate psychiatric screening for PD in *routine* CED care.

Material and methods

Study design

A cohort was formed of patients aged 18-70 years who presented with NCCP at the CED of VU-University Medical Center; Amsterdam, the Netherlands, between November 2012 and August 2013. Exclusion criteria included inadequate understanding of the Dutch language, an earlier CED visit within the study period, ongoing psychiatric treatment, and a likely or definitive somatic cause reported by the cardiologist. Eligible patients were asked to fill out a screening instrument consisting of the Hospital Anxiety and Depression Scale (HADS).

Measurement instruments

The HADS is a 14-item self-report questionnaire which is valid and reliable in populations with NCCP.[10] In accordance with earlier studies we used a cutoff score of 8 on either anxiety or depressive subscale, which yields a sensitivity of 98% for the presence of anxiety disorders.[1,10] Those scoring above cut-off were contacted by the Psychiatric Department to conduct the Composite International Diagnostic Interview (CIDI). The CIDI is a structured interview with good reliability and validity [11] and was administered by telephone within two weeks after CED discharge.

Implementation process

The study started with a pilot phase in which we performed barrier analyses in order to optimize the implementation methods.[12] The implementation process was adapted accordingly in three different ways. First, two CED nurses and a cardiology resident were made responsible for daily program evaluations. Second, administrative procedures were simplified. Finally, an experienced liaison psychiatrist (ADB) offered 1-hour training sessions to the CED-staff in effectively engaging patients with regard to psychiatric symptoms.

During the implementation phase CED-staff provided daily data on staff and patient-adherence. Staff-adherence was defined as proportion of eligible patients in whom screening was offered. Patient-adherence was defined as proportion of patients who filled out the screening tool if it was being offered. Monthly staff-adherence rates were fed back to the CED-staff in meetings by the researchers (NMB and AMB). New barriers to implementation and implementation goals were also identified and discussed in these meetings.

Analysis

Comparisons in gender, age and number of CED visits within the study period were made with Chi Squared statistics and one-way analysis of variance statistics (ANOVA). We compared patients with whom screening was initiated with those with whom it was not initiated as well as patients who refused screening with those who agreed to participate.

Ethical considerations

This study was performed in accordance with the principles of the Helsinki declaration and approval was obtained from the Ethics committee of the VU-University Medical Center.

Results

Feasibility

Staff adherence to the screening program was low, as only 60 out of 252 eligible patients (23.8%) were offered screening. A lack of time due to the primary task of providing *acute* cardiac care was reported most (88.0%) by the CED-staff.

Initial patient-adherence was higher, as 51 out of 60 patients (85.0%) agreed in screening. Most patients who refused participation saw no benefit in psychiatric screening in relation to their perceived life-threatening symptoms (n=6). However, in the second phase of screening, patient-adherence was low: twelve out of 24 patients (50%) refused administration of the CIDI, six of whom insisted on seeking psychiatric care with their General Practitioner (GP). See figure 1 for the flow-chart of inclusion into study.

During the course of this program we found no significant improvement in levels of adherence (data not shown). We were not able to resolve the barriers resulting in low adherence by offering assistance to administering the HADS nor by training staff to adequately address psychiatric problems.

Outcome

There were no differences in age, gender and number of CED visits between those who were offered screening and those who were not or between those who refused screening and those who participated. In 24 out of 51 patients HADS scores were above cut-off. Based on known prevalence numbers for PD of 12-41% in NCCP-patients [3,4] our cohort was estimated to include 38-130 PD patients. Ultimately, our screening program identified only 2 PD patients. Additionally, we classified Generalized Anxiety Disorder (n=1), Post-Traumatic Stress Disorder (n=1), Major Depressive Disorder (n=4), any Somatoform Disorder (n=4), and Alcohol Dependence (n=1).

Discussion and conclusion

The presence of heterogenic psychiatric disorders in patients with NCCP calls for a more personalized approach, instead of a screening program aimed at identifying those with PD. Screening refusal rates may be improved by approaching patients a couple of days after CED presentation, as done by Kuijpers et al.,[10] or by involvement from patients' GP, as some patients reported preferred consulting their GP. A limitation of this study was the sparse data-collection on reasons for low staff-adherence.

In a sample of CED patients with NCCP (n=252) we deemed screening for PD of limited feasibility when implemented in *routine* CED care. The main barriers were low staff-adherence and relatively high patient refusal rates.

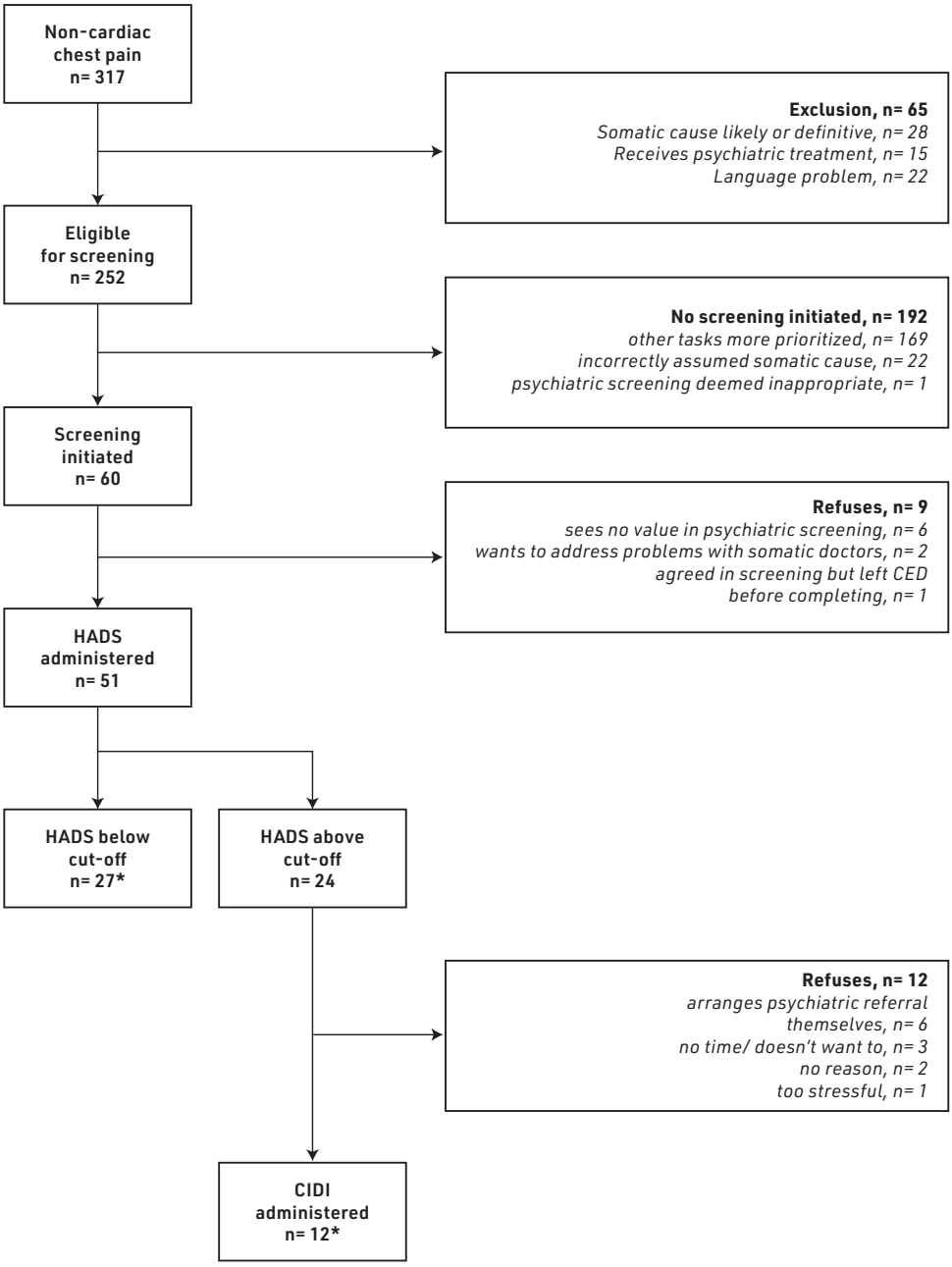


Figure 1. Flowchart of the screening program

* screening program completers

Future programs aimed at psychiatric screening in NCCP-patients should be performed after the acute phase, could benefit from GP involvement and should target a broad range of psychiatric disorders.

References

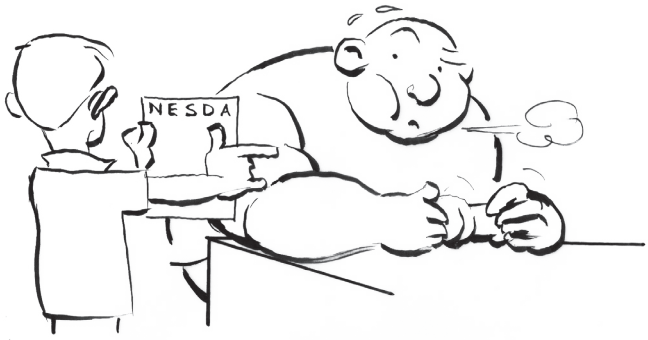
1. P.M.J.C. Kuijpers, A. Honig, E.J.L. Griez, S.H.J.G. Braat HJJW. Paniekstoornissen, pijn op de borst en palpitaties: een pilotonderzoek op een Nederlandse Eerste Harthulp. *Ned Tijdschr Geneeskd*. 2000:745-749.
2. Mayou RA, Bass C, Hart G, Tyndel S, Bryant B. Can clinical assessment of chest pain be made more therapeutic? *QJM*. 2000;93(12):805-811.
3. White KS, Raffa SD, Jakle KR, et al. Morbidity of DSM-IV Axis I disorders in patients with noncardiac chest pain: Psychiatric morbidity linked with increased pain and health care utilization. *J Consult Clin Psychol*. 2008;76(3):422-430.
4. Dammen T, Arnesen H, Ekeberg Ø, Friis S. Psychological factors, pain attribution and medical morbidity in chest-pain patients with and without coronary artery disease. *Gen Hosp Psychiatry*. 2004;26(6):463-469. doi:10.1016/j.genhosppsych.2004.08.004
5. Fleet RP, Dupuis G, Marchand A, Burelle D, Arsenault A, Beitman BD. Panic disorder in emergency department chest pain patients: Prevalence, comorbidity, suicidal ideation, and physician recognition. *Am J Med*. 1996;101(4):371-380.
6. Harvison KW, Woodruff-Borden J, Jeffery SE. Mismanagement of panic disorder in emergency departments: Contributors, costs, and implications for integrated models of care. *J Clin Psychol Med*. 2004;11:217-232.
7. Dammen T, Bringager CB, Arnesen H, Ekeberg Ø, Friis S. A 1-year follow-up study of chest-pain patients with and without panic disorder. *Gen Hosp Psychiatry*. 2006;28(6):516-524. doi:10.1016/j.genhosppsych.2006.08.005
8. Swinson RP, Soulios C, Cox BJ, Kuch K. Brief treatment of emergency room patients with panic attacks. *Am J Psychiatry*. 1992;149(7):944-946.
9. van Beek MHCT, Oude Voshaar RC, Beek a M, et al. A brief cognitive-behavioral intervention for treating depression and panic disorder in patients with noncardiac chest pain: a 24-week randomized controlled trial. *Depress Anxiety*. 2013;30(7):670-678. doi:10.1002/da.22106
10. Kuijpers PMJC, Denollet J, Lousberg R, Wellens HJJ, Crijns H, Honig A. Validity of the hospital anxiety and depression scale for use with patients with noncardiac chest pain. *Psychosomatics*. 2003;44(4):329-335. doi:10.1176/appi.psy.44.4.329
11. Haro JM, Arbabzadeh-Bouchez A, Brugha TS, et al. Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health Surveys. *Int J Methods Psychiatr Res*. 2006;15(4):167-180. doi:10.1002/mpr
12. Grol R, Wensing M. Effective implementation: a model. In: R. Grol, M. Wensing ME, ed. *Improving Patient Care: The Implementation of Change in Clinical Practice*. Edinburgh: Elsevier Butterworth- Heinemann; 2005:41-59.

Chapter 3.

Impact of Anxiety and/or Depressive Disorders and Chronic Somatic Diseases on disability and work impairment

Journal of Psychosomatic Research (2017)

Wicher A. Bokma
Neeltje. M. Batelaan
Anton J.L.M. van Balkom
Brenda W.J.H. Penninx



Abstract

Objective

Anxiety and/or Depressive Disorders (ADDs) and Chronic Somatic Diseases (CSDs) are associated with substantial levels of health-related disability and work impairment. However, it is unclear whether comorbid ADDs and CSDs additively affect functional outcomes. This paper examines the impact of ADDs, CSDs, and their comorbidity on disability, work absenteeism and presenteeism.

Methods

Baseline data from the Netherlands Study of Depression and Anxiety (n=2,371) were used. We assessed presence of current ADDs (using psychiatric interviews, CIDI) and presence of self-reported CSDs. Outcome measures were disability scores (WHO-DAS II questionnaire, overall and domain-specific), work absenteeism (≤2 weeks and >2 weeks; TiC-P) and presenteeism (reduced and impaired work performance; TiC-P). We conducted multivariate regression analyses adjusted for socio-demographics.

Results

Both ADDs and CSDs significantly and independently impact total disability, but the impact was substantially larger for ADDs (main effect unstandardized β=20.1, p<.001) than for CSDs (main effect unstandardized β=3.88, p<.001). There was a positive interaction between ADDs and CSDs on disability (unstandardized β interaction=4.06, p=.004). Although CSDs also induce absenteeism (OR for extended absenteeism=1.42, p=.015) and presenteeism (OR for impaired work performance=1.42, p=.013), associations with ADDs were stronger (OR for extended absenteeism=6.64, p<.001; OR for impaired work performance=7.51, p<.001).

Conclusion

Both CSDs and ADDs cause substantial disability, work absenteeism and presenteeism, but the impact of ADDs far exceeds that of CSDs. CSDs and ADDs interact synergistically on disability, thereby bolstering the current view that patients with physical mental comorbidity (PM-comorbidity) form a severe subgroup with an unfavourable prognosis.

Introduction

Disability and work impairment are important indicators of poor health, from both a societal and a clinical perspective.[1,2] Those with Anxiety and/or Depressive Disorders (ADDs; either Anxiety Disorders or Depressive Disorders) or Chronic Somatic Diseases (CSDs) are known to suffer from many years lived with disability (YLD).[3–5] Globally, YLD are highest in low back pain, with other CSDs such as iron deficiency anaemia, other musculoskeletal disorders, lung disease, migraine, and diabetes among the top ten most disabling diseases. Among the ADDs, Depressive Disorders ranked second and Anxiety Disorders ranked sixth.[5] Those with CSDs and ADDs also suffer from substantial levels of work impairment.[1,3,6] However, CSDs and ADDs were found to frequently co-occur: a phenomenon referred to as physical mental comorbidity (PM-comorbidity).[3,6–8] Among those with ADDs, higher incidences of lung disease, diabetes, obesity, cardiovascular diseases, hypertension, stroke, Alzheimer's disease, peptic ulcers, symptoms of Irritable Bowel Syndrome, and osteoarthritis have been found.[7,9] Likewise, higher incidences of ADDs were found among those with lung diseases, hypertension, allergies, peptic ulcers, autoimmune disease, thyroid disease, chronic back problems, osteoarthritis, and migraine.[6,8,10,11] Therefore, it is evident that a wide variety of CSDs form PM-comorbidity with ADDs. The current literature on PM-comorbidity suggests that it forms a relevant subgroup characterized by a worse prognosis with regard to several clinical outcome measures, including functional outcomes [4,6,11] and less favourable CSD-related treatment response.[12–14] Despite the known separate impact of CSDs and ADDs on disability and work impairment, the high prevalence of PM-comorbidity, and its associations with unfavourable health-related outcomes, little is known of the effect of PM-comorbidity on disability and work impairment.

A number of studies on disability found that comorbidity with ADDs increased disability associated with CSDs.[3,15] Both Armenian (1998) and Stein (2006) assessed interaction effects between CSDs and ADDs on disability, but whereas Armenian found an interaction effect, Stein did not.[15,16] However, these studies included a limited number of CSDs, and separate interaction effects for specific CSDs or specific disability domains were not reported. With regard to work impairment, a number of studies found increased work impairment in those with PM-comorbidity, compared to those with either ADDs or CSDs alone.[6,11,17,18] Kessler et al. [6] and Buist-Bouwman et al. [11] found interaction effects between ADDs and CSDs on work impairment. However,

the latter studies only included a limited number of CSDs and assessed absenteeism (absence from work due to health issues) but not presenteeism (presence at work while hindered by health issues), while presenteeism is regarded a highly prevalent and costly form of work impairment.[19,20] These inconclusive findings warrant further research to clarify the nature of associations of CSDs, ADDs, and PM-comorbidity with regard to disability, work absenteeism and presenteeism.

Aims of the study

We aim to expand on the current literature by studying severity of disability, work absenteeism and presenteeism associated with Anxiety and/or Depressive Disorders (ADDs), Chronic Somatic Diseases (CSDs) and their comorbidity in a wide range of CSDs. This paper examines the relative separate effects of CSDs and ADDs on total disability, disability domains, work absenteeism and presenteeism. In addition to the separate effects, we will assess whether synergistic effects (i.e. positive interaction effects) between CSDs and ADDs exist. We expect ADDs and CSDs to have substantial separate main effects on disability and work impairment, and expect a positive interaction effect in those with PM-comorbidity.

Method

Design and sample

Respondents were derived from the Netherlands Study of Depression and Anxiety (NESDA), an ongoing cohort study consisting of 2,981 respondents (aged 18-65) at baseline. Since the aim of NESDA is to gain insight into the long-term course and consequences of Anxiety and Depressive Disorders, those with Anxiety Disorders or Depressive Disorders were oversampled. NESDA recruitment took place in three settings: community, primary care, and specialized mental health care, in order to represent all developmental stages of ADDs. Exclusion criteria included a primary diagnosis of other psychiatric disorders, such as bipolar, obsessive compulsive, substance use or psychotic disorders and insufficient command of the Dutch language. Baseline assessments were conducted between 2004 and 2007 and included a structured diagnostic psychiatric interview. A full description of the NESDA study design is available elsewhere.[21] The Ethical Committee of participating universities approved of the study protocol and all respondents provided written informed consent. The current study uses the baseline data and included persons with presence of current (i.e. six-month) ADDs (*n*=1,737),

and controls without current and lifetime presence of ADDs (*n*=634). We excluded 610 respondents due to presence of lifetime, but not current, diagnoses of ADDs.

Anxiety and/or Depressive Disorders (ADDs)

ADDs were defined as presence of either a Depressive Disorder (Depressive or Dysthymic Disorder) or an Anxiety Disorder (Generalized Anxiety Disorder, Social Phobia, Panic Disorder with or without Agoraphobia). We assessed Depressive and Anxiety Disorders combined since both groups of disorders are associated with increased disability [3-5] and comorbidity levels between these disorders are known to be high in other studies [22] but also in our own study.[23] Presence of ADDs was assessed using the Composite International Diagnostic Interview (CIDI, version 2.1), which classifies diagnoses according to DSM-IV criteria.[24,25] The CIDI has good overall reliability and validity and is frequently used worldwide.[26] The structured CIDI interviews were conducted by highly trained staff.

Chronic Somatic Diseases (CSDs)

A 21-item face-to-face interview was used to assess presence of CSDs.[21] This instrument was used previously in large-scale population-based cohort studies.[6,11,27] Respondents were asked for presence of 30 CSDs and were able to report any additional CSDs they may have. Individual CSDs were deemed present when respondents reported monitoring or receiving prescription medication by a General Practitioner or a medical specialist for that CSD. Following earlier research,[28] we clustered separate CSDs into seven disease categories: respiratory, cardio-metabolic, musculoskeletal, gastrointestinal, neurological, endocrine and cancer. We used presence of any CSD and presence of each CSD category as outcome measures.

Disability

Disability during the previous 30 days was assessed using the WHO-Disability Assessment Schedule (WHO-DAS II), a 36-item self-report questionnaire.[29] It measures disability in six domains: cognition (six items, Cronbach's α=.92 in our sample), mobility (five items, α=.91), self-care (four items, α=.84), interpersonal interactions (five items, α=.88), household activities (five items, α=.95), and participation in society (eight items, α =.92) on a 5-point Likert scale with item scores ranging from 0 (no difficulties) to 4 (extreme difficulties/cannot do). We excluded four items concerning work-related disability, as a substantial proportion of our sample (n=905) was neither currently employed for at least eight hours a week nor attending education.

Domain scores were calculated by adding all domain item scores and a total disability score was calculated by adding all 32 item scores. There were 49 respondents with missing data on WHO-DAS data; we replaced missing scores with mean scale values of total scale scores. Domain and total scores were standardized to derive scores ranging from 0 to 100, with higher scores indicating higher levels of disability.

Work impairment

Work impairment was analysed within a subsample of employed participants, which we defined as having a paid job for at least eight hours a week divided over more than one day a week (*n*=1,466), thereby excluding 905 respondents who were not employed, or who were employed for less than eight hours a week. We excluded another four respondents due to missing values on work impairment data, which yielded a sample of *n*=1,462. We used the Trimbos/iMTA questionnaire for Costs Associated with Psychiatric Illness (TiC-P) to assess two aspects of work impairment: absenteeism and presenteeism.[30] Absenteeism was calculated by dividing the total number of hours that respondents were absent from work during the previous six months by the number of hours that respondents were supposed to work per week. Work absenteeism is measured in weeks and ranges from 0 to 26 weeks. Presenteeism is defined as the number of workweeks in which quality of work was reduced due to health issues, multiplied by a self-reported proportional score for severity of work quality reduction.[31] Presenteeism scores ranged from 0 to 26. As absenteeism and presenteeism data did not meet normality assumptions, we categorized these into 'no absenteeism', 'short absenteeism' (≤2 weeks), and 'extended absenteeism' (>2 weeks); and 'no presenteeism' (score=0), 'reduced work performance' (0< highest quartile) and 'impaired work performance' (>highest quartile), as done previously.[19,31]

Statistical analyses

To compare baseline characteristics of ADD patients to controls, two-tailed Pearson's Chi Squared statistics were used for categorical variables and independent samples t-tests were used for continuous variables.

To compare presence of CSDs in ADD patients versus controls, we performed logistic regression analyses, which yielded Odds Ratios (OR) for ADD patients for having any CSD and for having each of the seven CSD categories.

To test the impact of ADDs, CSDs, and their interaction on disability total and domain scores, we performed multivariate linear regression analyses using presence of ADDs (0=no, 1=yes), CSDs (0=no, 1=yes), and their interaction term (ADD*CSD) as independent variables, and the WHO-DAS II total and domain scores as dependent variables. In this model, interactions were tested on an additive scale, with a significant positive interaction effect implying that the effect of comorbid CSDs and ADDs on disability scores is greater than the sum of separate effects for CSDs and ADDs: synergy. In addition, impact on disability among different CSD categories was examined by repeating these analyses on total disability scores using seven CSD category variables instead of the overall CSDs variable.

We tested the impact of ADDs, CSDs, and PM-comorbidity on work absenteeism and presenteeism by performing multinomial logistic regression analyses with categorized absenteeism and presenteeism as dependent variables. First, we determined main effects for ADDs and CSDs; second, we divided our sample into those with purely physical disorders (presence of any CSD but absence of ADDs), those with purely mental disorders (presence of ADD but absence of CSDs) and those with PM-comorbidity (presence of CSD and ADD).

Finally, we performed several sensitivity analyses in order to check for possible bias. First, we considered the possibility that antidepressant side effects caused respondents to report presence of CSDs by conducting logistic regression analyses for presence of CSDs in those with ADDs, adjusted for antidepressant use. Second, we checked the possibility that ADD patients overreported presence of CSDs by performing logistic regression analyses for presence of CSDs with a more stringent criterion of CSDs based on medication use. Third, we repeated linear regression analyses on total disability to assess whether associations with disability are different for those with more severe CSDs or ADDs, compared to those with less severe CSDs or ADDs. A complete list of methods used have been added to the supplement. We used an α-value of .05 for all our analyses. All regression analyses were adjusted for covariates age (years), sex, and education (years). Statistical analyses were performed using SPSS Statistics, version 20 (IBM Corp., USA).

Results

Sample

Table I shows the baseline characteristics of the total sample (*n*=2,371) and the employed subsample (*n*=1,462). ADD patients were more likely to be female, less educated, and to have higher levels of disability and work impairment as compared to controls. Table II shows that CSDs were more often present in ADD patients (42.8%) than in controls (35.6%) (p=.002). Odds Ratio for presence of any CSD was 1.34 (95%-CI:1.09-1.64). Especially the odds for having gastrointestinal disease were significantly raised in ADD patients (Table II). Figure 1 shows the unadjusted mean standardized total disability scores stratified for presence of CSDs and ADDs.

Impact of CSDs, ADDs, and PM-comorbidity on disability

Multivariate regression analysis on total disability showed that the main effect for CSDs (unstandardized β=3.88, p<.001) was surpassed by that of ADDs (unstandardized β=20.1, p<.001), see Table III (Model 1). Both ADDs and CSDs negatively influenced all disability domains. The main effect of ADDs was greatest in the domains of 'participation in society', 'cognition', 'interpersonal interactions', and 'household activities', whereas the main effect of CSDs was greatest in the domains of 'mobility', 'self-care' and 'participation in society'. Nevertheless, on all domains, the regression coefficients of ADDs were 3-5 times larger than those of CSDs. The interaction term for CSD*ADD on total disability was positive and statistically significant (unstandardized β=4.06, p=.004, Table III, Model 2). This suggests that, on average, those with PM-comorbidity suffer from disability levels that exceed the sum of disability associated with CSDs and ADDs. This translates to an additional 4-5 points on the total WHO-DAS II score. The interaction term CSD*ADD was positive and statistically significant in three domains: 'mobility', 'household activities' and 'self-care'. The magnitude of impact of CSDs on disability was not driven by specific CSD categories. Main effects on total disability score for different CSD categories ranged from 0.70 to 4.65 (data not shown).

Impact of CSDs, ADDs, and PM-comorbidity on work impairment

Table IV (Model 1) shows that presence of ADDs was associated with worse absenteeism and presenteeism outcomes in an employed subsample (n=1,462): short absenteeism (OR 2.88, 95%CI: 2.16-3.84), extended absenteeism (OR 6.64, 95%CI: 4.69-9.40), reduced work performance (OR 1.83, 95%CI: 1.38-2.43), and impaired work performance (OR 7.51

Table I. Baseline characteristics for respondents with current (six month) Anxiety and/or Depressive Disorders[b] (ADDs) compared to controls without lifetime diagnosis of ADDs[b] (*n*=2,371 for total sample, *n*=1,462 for employed subsample).

	No lifetime Anxiety and/or Depressive Disorder[b]		**Current Anxiety and/or Depressive Disorder[b]**				
	(*n*=634)		**(*n*=1,737)**		t	χ^2	p
Socio-demographics							
Gender, female	61.7%		67.0%			5.74	**.017**
Age, mean ±SD	41.1	±14.7	41.3	±12.4	-0.45		.681
Education (in years), mean ±SD	12.8	±3.2	11.8	±3.3	7.01		**<.001**
Somatic variables, mean ±SD							
Number of chronic somatic disease categories	0.47	±0.74	0.63	±0.86	-4.06		**<.001**
Disability variables, mean ±SD a							
WHO-DAS II, total score	7.8	±9.3	29.0	±16.4	-30.7		**<.001**
Domains:							
Cognition	8.4	±11.5	32.4	±20.5	-27.8		**<.001**
Mobility	5.1	±12.0	19.0	±21.6	-15.3		**<.001**
Self-care	3.0	±8.5	14.7	±17.9	-15.9		**<.001**
Interpersonal interactions	9.5	±13.7	35.3	±23.0	-26.5		**<.001**
Household activities	12.6	±18.3	39.5	±27.0	-23.3		**<.001**
Participation in society	8.2	±11.3	33.5	±20.1	-30.0		**<.001**
	(*n*=424)		**(*n*=1,038)**				
Work functioning variables							
Absenteeism, last six months[c]							
No absenteeism	67.2%		32.5%			163.9	**<.001**
Short absenteeism (≤2 weeks)	21.9%		30.3%				
Extended absenteeism (>2 weeks)	10.8%		37.3%				
Presenteeism (score), last six months[d]							
No presenteeism (0)	67.5%		39.5%			132.2	**<.001**
Reduced work performance (0 < highest quartile)	24.5%		25.0%				
Impaired work performance (>highest quartile)	8.0%		35.5%				

[a] a higher score indicates higher severity of disability, standardized scores (range 0-100).
[b] Anxiety and/or Depressive Disorders were defined as presence of either an Anxiety Disorder (Social Phobia, Generalized Anxiety Disorder, Panic Disorder with or without Agoraphobia) or a Depressive Disorder (Dysthymic of Major Depressive Disorder).
[c] absenteeism is represented by number of weeks absent from work during the last six months.
[d] presenteeism is represented by number of workweeks in which quality of work was reduced due to health issues, multiplied by a self-reported proportional score for severity of work quality reduction.

Table II. Odds of having presence of Chronic Somatic Diseases (CSDs) for Anxiety and/or Depressive Disorder[a] (ADD) patients in comparison with controls (*n*=2,371).

		No lifetime Anxiety and/or Depressive Disorder[a] (*n*=634)	**Current Anxiety and/or Depressive Disorder[a] (*n*=1,737)**			
		%	%	OR	(95% CI)	p
Any chronic somatic disease	All those listed below	35.6	42.8	1.34	(1.09-1.64)	.002
Respiratory	Asthma, chronic bronchitis, pulmonary emphysema, other lung diseases	6.8	10.0	1.39	(0.98-1.98)	.067
Cardio-metabolic	Hypertension, angina pectoris, history of cardiac disease, stroke, diabetes, vascular abnormalities	16.6	16.2	1.05	(0.79-1.39)	.745
Musculoskeletal	Osteoarthritis, rheumatoid arthritis, systemic lupus erythematodes, fibromyalgia, RSI, congenital skeletal deformation	8.2	10.8	1.33	(0.95-1.85)	.101
Gastrointestinal	Ulcer, irritable bowel syndrome, Crohn's disease, colitis ulcerosa, diverticulitis, liver cirrhosis, hepatitis, constipation, oesophageal disease, gastric sphincter dysfunction, other gastrointestinal disease	3.9	12.2	**3.29**	**(2.15-5.05)**	**<.001**
Neurological	Migraine, epilepsy, multiple sclerosis, peripheral neuropathy, hernia	2.2	3.7	1.78	(0.98-3.22)	.057
Endocrine	Thyroid dysfunction	2.8	2.9	1.12	(0.64-1.96)	.696
Cancer	Throat, thyroid, lymphoid, lung, oesophagus, bowel, stomach, liver, uterus, cervix, ovary, bladder, testicle, prostate, skin, brain, blood	6.3	6.8	1.10	(0.75-1.61)	.633

All ORs were adjusted for gender, age, and education.
[a] Anxiety and/or Depressive Disorders were defined as presence of either an Anxiety Disorder (Social Phobia, Generalized Anxiety Disorder, Panic Disorder with or without Agoraphobia) or a Depressive Disorder (Dysthymic of Major Depressive Disorder)

95%CI: 5.11-11.1). Presence of CSDs was solely associated with extended absenteeism (OR 1.42, 95%CI: 1.07-1.88) and impaired work performance (OR 1.42, 95%CI: 1.08-1.87). Table IV (Model 2) shows that odds for each form of absenteeism and presenteeism are highest in those with PM-comorbidity, followed by those with ADDs (without CSDs) and those with CSDs (without ADDs). Effects of separate CSD categories on work impairment outcomes were less clear. Respiratory and musculoskeletal diseases were associated with increased odds for extended absenteeism (OR 1.65, 95%CI 1.01-2.69 for respiratory diseases and OR 2.65, 95%CI 1.60-4.41 for musculoskeletal diseases), whereas cardio-metabolic diseases were associated with decreased odds for extended absenteeism (OR 0.64, 95%CI 0.43-0.97). No associations between separate CSD categories and presenteeism existed (data not shown).

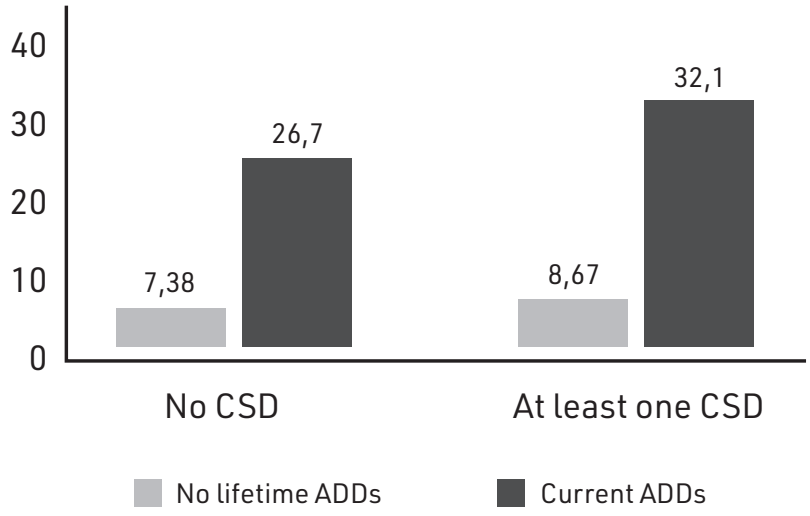


Figure 1. Mean levels of disability (standardized WHO-DAS II total score) for those with any self-reported Chronic Somatic Diseases (CSDs), those with Anxiety and/or Depressive Disorders (ADDs) and those with PM-comorbidity.

Sensitivity analyses

Sensitivity analyses showed that antidepressant use was not accountable for differences in presence of CSDs in an antidepressant-adjusted model (see supplement). Furthermore, there was no evidence for a self-report bias of CSDs in ADD patients, as associations with CSDs in ADD patients were

comparable when applying more stringent CSD definitions. Additionally, the positive interaction effects with ADDs on total disability were comparable for those with only one CSD and those with multiple CSDs. Finally, although the total disability scores are highest in those with comorbid Anxiety and Depressive Disorders, the positive interaction effects with CSD status were comparable for those with pure Anxiety or Depression compared to those with comorbid Anxiety and Depressive Disorders. All sensitivity analyses can be found in the supplement.

Table III. Adjusted unstandardized regression coefficients for the association between Chronic Somatic Diseases (CSDs) and Anxiety and/or Depressive Disorders[b] (ADDs) on total disability and on disability domains (*n*=2,371).

Level of disability (WHO-DAS II)[a]		**Total score (32 items)**	**Cognition**	**Mobility**	**Self-care**	**Interpersonal interactions**	**Household activities**	**Participation in society**
	n	*β*	*β*	*β*	*β*	*β*	*β*	*β*
Model 1, main effects								
Constant		**12.4****	**18.5****	**8.84****	**14.0****	**12.7****	**6.00**	**14.1****
Any CSD	969	**3.88****	**2.32****	**6.05****	**3.72****	**2.87****	**4.93****	**4.05****
ADDs[b]	1,737	**20.1****	**23.0****	**12.0****	**10.7****	**25.3****	**25.6****	**24.3****
Model 2, main effects and interaction effect								
Constant		**13.1****	**18.8****	**10.3****	**14.7****	**13.4****	**7.29***	**14.4****
Any CSD	969	0.80	1.15	0.22	1.12	0.20	-0.29	2.68
ADDs[b]	1,737	**18.6****	**22.4****	**9.13****	**9.38****	**24.0****	**23.1****	**23.6****
Interaction term CSD*ADD[c]	743	**4.06****	1.53	**7.68****	**3.42***	3.51	**6.86****	1.81

These regression models were controlled for socio-demographics (gender, education, and age), β are unstandardized regression coefficients.

[a] WHO-DAS total and domain scores were standardized to values 0-100, a higher score indicates higher severity of disability.

*: p<.05, **: p<.01.

[b] Anxiety and/or Depressive Disorders were defined as presence of either an Anxiety Disorder (Social Phobia, Generalized Anxiety Disorder, Panic Disorder with or without Agoraphobia) or a Depressive Disorder (Dysthymic of Major Depressive Disorder)

[c] in this model, interactions are tested on an additive scale, a significant positive interaction effect (for instance: β *interaction*=4.06) implies that the effect of comorbid CSD and ADD on disability score is larger than the sum of separate effects for CSD and ADD (i.e. synergistic effect modification).

Table IV. Multinomial regression coefficients for impact of Chronic Somatic Diseases (CSDs) and current Anxiety and/or Depressive Disorders (ADDs) on absenteeism and presenteeism in employed respondents (*n*=1,462).

		Absenteeism[a]				Presenteeism[b]			
		Short absenteeism (<2 weeks) vs. no absenteeism (ref)		Extended absenteeism (>2 weeks) vs. no absenteeism (ref)		Reduced work performance vs. no presenteeism (ref)		Impaired work performance vs. no presenteeism (ref)	
	n	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)
Model 1, main effects									
Any CSD	536	1.07	(0.80-1.42)	**1.42**	**(1.07-1.88)***	1.06	(0.80-1.41)	**1.42**	**(1.08-1.87)***
ADDs[c]	1,038	**2.88**	**(2.16-3.84)****	**6.64**	**(4.69-9.40)****	**1.83**	**(1.38-2.43)****	**7.51**	**(5.11-11.1)****
Model 2, main effects									
Controls *No CSDs, no ADDs*	281	1.00	(ref)	1.00	(ref)	1.00	(ref)	1.00	(ref)
Purely physical *CSDs present, no ADDs*	143	1.03	(0.61-1.74)	1.44	(0.75-2.75)	0.99	(0.59-1.67)	**3.22**	**(1.54-6.73)****
Purely mental *no CSDs, ADD present*	645	**2.84**	**(2.00-4.02)****	**6.68**	**(4.27-10.5)****	**1.81**	**(1.29-2.53)****	**11.6**	**(6.55-20.7)****
Physical mental (PM-)comorbidity *CSD present, ADD present*	393	**3.07**	**(2.05-4.59)****	**9.50**	**(5.88-15.3)****	**1.93**	**(1.30- 2.85)****	**14.9**	**(8.17-27.1)****

These regression models were controlled for socio-demographics (gender, education, age).
[a] absenteeism is represented by number of weeks absent from work during the last six months.
[b] presenteeism is represented by number of workweeks in which quality of work was reduced due to health issues, multiplied by a self-reported proportional score for severity of work quality reduction.
*: p<.05, **: p<.01.
[c] Anxiety and/or Depressive Disorders were defined as presence of either an Anxiety Disorder (Social Phobia, Generalized Anxiety Disorder, Panic Disorder with or without Agoraphobia) or a Depressive Disorder (Dysthymic of Major Depressive Disorder)

Discussion

This paper aims to examine the impact of ADDs, CSDs, and PM-comorbidity on disability and work impairment. First, we found that the relative impact of ADDs far exceeded that of CSDs on total disability, disability domains, work absenteeism, and presenteeism. Second, a synergistic effect existed between CSDs and ADDs with regard to disability; those with PM-comorbidity were burdened with levels of disability that were more than what would be expected on the basis of separate CSD and ADD effects. Likewise, work absenteeism and presenteeism were affected the most in those with PM-

comorbidity. This is the first paper to establish this association for work presenteeism. Third, this synergistic effect of CSDs and ADDs was present in a number of disability domains, but not dependent on specific separate CSD categories.

Levels of disability and work impairment were substantial in our sample. Our finding that mean disability and odds of worse work impairment outcomes are higher in those with ADDs than those with CSDs is in accordance with earlier research, and underscores the high burden of disease associated with ADDs.[3,15] This largely corresponds with findings elsewhere, although Buist-Bouwman (2006) reports mobility to be affected more substantially by CSDs compared to ADDs.[3,5] This difference could be because Buist-Bouwman et al. included only arthritis and heart disease. Moreover, sensitivity analyses did not show an increase of effect of CSDs on disability or work impairment when applying a more stringent criterion for presence of CSDs (see supplement), thereby making it unlikely that the higher impact of ADDs is due to misclassification of CSDs.

Most notably, we were able to replicate a positive interaction effect on disability found earlier by Armenian (1998).[16] Whereas Armenian et al. did not quantify the magnitude of this synergistic effect, we did: those with PM-comorbidity suffered from disability levels which exceeded the sum of disability associated with CSDs and ADDs by 4-5 points on the total (unstandardized) WHO-DAS scale. Our analyses showed that those with PM-comorbidity also have the highest rates of work impairment. Furthermore, in addition to work absenteeism, we found increased rates of reduced and impaired work performance in those with PM-comorbidity. This could indicate that loss of work productivity in those with PM-comorbidity adds to the already high societal costs of PM-comorbidity.[20,31]

These findings indicate that those with PM-comorbidity should be regarded as a distinctly identifiable and highly burdened subgroup. Moreover, there is growing evidence in current literature that standard treatment is less effective in those with PM-comorbidity. For instance, Anxiety and/ Depressive Disorder (ADD) patients are known to have lower levels of adherence to somatic medication [12] and require different psychosocial approaches.[32] Antidepressant medication, frequently used in ADDs, may induce metabolic syndrome, which could cause additional disability by causing cardiovascular events.[14] Additionally, some medications regularly used in somatic medicine may worsen psychiatric functioning.[33] As a result,

some authors argue that tailored rehabilitation programmes should be employed in order to effectively deal with disability and work impairment in those with PM-comorbidity.[32,34] This is in line with the theoretical concepts of applying clinical staging models in psychiatry, for which an increased amount of interest has been shown in recent years.[35,36] Currently, clinical staging models for psychiatric disorders are being developed. Our findings implicate that presence of PM-comorbidity should be incorporated in future clinical staging models for psychiatric disorders due to the evident effects on functional outcomes and prognosis.

In our sample, interaction effects were statistically significant for the more physically oriented domains of disability, i.e. mobility, self-care, and household activities. This implicates that, by means of a synergistic effect with CSDs, ADDs have a bigger impact on physical outcome measures than is apparent from their individual coefficient. We did not find separate CSD category effects that exceeded the overall CSD effect, therefore increased disability and work impairment rates were not driven by specific CSD categories.

The main strength of this study is the extensive way in which the interaction effects between CSDs and ADDs were examined on various outcome measures in a large sample. Furthermore, the generalizability of findings from this study was increased by examining a wide range of CSDs as opposed to a limited number. The present study had several limitations. First, for assessing presence of CSDs we relied on self-report evaluation, which is accompanied by a risk of self-report bias. Still, good concordance between self-report and diagnoses by a medical doctor were found [28,37] and our sensitivity analyses ruled out the possibilities of self-report bias of CSDs in those with ADDs, as associations remained comparable when applying more stringent CSD criteria (see supplement), thereby validating the use of our less stringent CSD classification. Moreover, we ruled out the possibility that presence of CSDs in ADD patients was due to antidepressant side effects. Second, as our sample is relatively young, prevalence of serious CSDs is bound to be relatively low as prevalence of serious CSDs increases dramatically in ageing populations.[38] Third, by analysing Anxiety Disorders and Depressive Disorders simultaneously, our population became more heterogeneous. Post hoc analyses showed that although severity of disability and work impairment was dependent on psychiatric diagnosis (comorbidity between Anxiety and Depressive Disorders induced the most disability and work impairment, followed by Depressive Disorders and Anxiety Disorders);

interaction effects showed comparable results among different psychiatric diagnoses (supplementary Table VI) and among those with single or multiple CSDs (supplementary Table V). This is a strong argument for examining both disorder groups together with regard to disability, work absenteeism and presenteeism.

To conclude, PM-comorbidity is associated with additional disability and high rates of work absenteeism and presenteeism and should be regarded as a clinically distinct subgroup of patients with a worse prognosis. Given the synergistic effect on disability and the high rates of work impairment in those with PM-comorbidity, more research into the efficacy of targeted interventions for those with PM-comorbidity is needed to improve functioning in this group.

References

1. de Graaf R, Tuithof M, van Dorsselaer S, ten Have M. Comparing the effects on work performance of mental and physical disorders. *Soc Psychiatry Psychiatr Epidemiol*. 2012;47(11):1873-1883. doi:10.1007/s00127-012-0496-7
2. World Health Organization, WHO. *World Report on Disability 2011*. Geneva: World Health Organisation; 2011. doi:10.1136/ip.2007.018143
3. Buist-Bouwman MA, De Graaf R, Vollebergh WAM, Alonso J, Bruffaerts R, Ormel J. Functional disability of mental disorders and comparison with physical disorders: a study among the general population of six European countries. *Acta Psychiatr Scand*. 2006;113(6):492-500. doi:10.1111/j.1600-0447.2005.00684.x
4. Ormel J, VonKorff M, Ustun TB, Pini S, Korten A, Oldehinkel T. Common Mental Disorders and Disability Across Cultures. *JAMA*. 1994;272(22):1741-1748. doi:10.1001/jama.1994.03520220035028
5. Vos T, Flaxman AD, Naghavi M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012;380(9859):2163-2196. doi:10.1016/S0140-6736(12)61729-2
6. Kessler RC, Ormel J, Demler O, Stang PE. Comorbid mental disorders account for the role impairment of commonly occurring chronic physical disorders: results from the National Comorbidity Survey. *J Occup Environ Med*. 2003;45(12):1257-1266. doi:10.1097/01.jom.0000100000.70011.bb
7. Wells KB, Golding JM, Burnam MA. Chronic medical conditions in a sample of the general population with anxiety, affective, and substance use disorders. *Am J Psychiatry*. 1989;146(11):1440-1446. doi:10.1176/ajp.146.11.1440
8. Sareen J, Cox BJ, Clara I, Asmundson GJG. The relationship between anxiety disorders and physical disorders in the U.S. National Comorbidity Survey. *Depress Anxiety*. 2005;21(4):193-202. doi:10.1002/da.20072
9. Penninx BWJH, Milaneschi Y, Lamers F, Vogelzangs N. Understanding the somatic consequences of depression: biological mechanisms and the role of depression symptom profile. *BMC Med*. 2013;11:129. doi:10.1186/1741-7015-11-129
10. Roy-Byrne PP, Davidson KW, Kessler RC, et al. Anxiety disorders and comorbid medical illness. *Gen Hosp Psychiatry*. 2008;30(3):208-225. doi:10.1016/j.genhosppsych.2007.12.006
11. Buist-Bouwman MA, de Graaf R, Vollebergh WAM, Ormel J. Comorbidity of physical and mental disorders and the effect on work-loss days. *Acta Psychiatr Scand*. 2005;111(6):436-443. doi:10.1111/j.1600-0447.2005.00513.x

12. Gentil L, Vasiliadis HM, Préville M, et al. Association between depressive and anxiety disorders and adherence to antihypertensive medication in community-living elderly adults. *J Am Geriatr Soc*. 2012;60(12):2297-2301. doi:10.1111/j.1532-5415.2012.04239.x
13. Rosenthal MH. The challenge of comorbid disorders in patients with depression. *J Am Osteopath Assoc*. 2003;103(8):10-15.
14. McIntyre RS, Alsuwaidan M, Goldstein BI, et al. The Canadian Network for Mood and Anxiety Treatments (CANMAT) task force recommendations for the management of patients with mood disorders and comorbid metabolic disorders. *Ann Clin Psychiatry*. 2012;24(1):69-81.
15. Stein MB, Cox BJ, Afifi TO, Belik SL, Sareen J. Does co-morbid depressive illness magnify the impact of chronic physical illness? A population-based perspective. *Psychol Med*. 2006;36(5):587-596. doi:10.1017/S0033291706007239
16. Armenian HK, Pratt LA, Gallo J, Eaton WW. Psychopathology as a predictor of disability: a population-based follow-up study in Baltimore, Maryland. *Am J Epidemiol*. 1998;148(3):269-275.
17. Ervasti J, Vahtera J, Pentti J, et al. The Role of Psychiatric, Cardiometabolic, and Musculoskeletal Comorbidity in the Recurrence of Depression-Related Work Disability. *Depress Anxiety*. 2014;31(9):796-803. doi:10.1002/da.22286
18. Ervasti J, Vahtera J, Pentti J, et al. Return to work after depression-related absence by employees with and without other health conditions: a cohort study. *Psychosom Med*. 2015;77(2):126-135. doi:10.1097/PSY.0000000000000138
19. Uegaki K, de Bruijne MC, Anema JR, van der Beek AJ, van Tulder MW, van Mechelen W. Consensus-based findings and recommendations for estimating the costs of health-related productivity loss from a company's perspective. *Scand J Work Environ Health*. 2007;33(2):122-130. doi:10.5271/sjweh.1115
20. Bergström G, Bodin L, Hagberg J, Aronsson G, Josephson M. Sickness presenteeism today, sickness absenteeism tomorrow? A prospective study on sickness presenteeism and future sickness absenteeism. *J Occup Environ Med*. 2009;51(6):629-638. doi:10.1097/JOM.0b013e3181a8281b
21. Penninx BWJH, Beekman ATF, Smit JH, et al. The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *Int J Methods Psychiatr Res*. 2008;17(3):121-140. doi:10.1002/mpr.256
22. Kessler RC, Chiu WT, Demler O, Walters EE, Merikangas KR. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(June 2005):617-627. doi:10.1001/archpsyc.62.6.617

23. Lamers F, van Oppen P, Comijs HC, et al. Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). *J Clin Psychiatry*. 2011;72(3):341-348. doi:10.4088/JCP.10m06176blu
24. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR Fourth Edition (Text Revision)*.; 2000.
25. World Health Organization. *World Health Organization, Composite International Diagnostic Interview (CIDI), Core Version 2.1*. Geneva; 1998.
26. Wittchen H. Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): a critical review. *J Psychiatr Res*. 1994;28:57-84.
27. Bijl R V., van Zessen G, Ravelli A, de Rijk C, Langendoen Y. The Netherlands Mental Health Survey and Incidence Study (NEMESIS): objectives and design. *Soc Psychiatry Psychiatr Epidemiol*. 1998;33(12):581-586. doi:10.1007/s001270050097
28. Gerrits MM, van Oppen P, van Marwijk HW, van der Horst H, Penninx BW. The impact of chronic somatic diseases on the course of depressive and anxiety disorders. *Psychother Psychosom*. 2013;82(1):64-66. doi:10.1159/000338636
29. Chwastiak LA, Von Korff M. Disability in depression and back pain: Evaluation of the World Health Organization Disability Assessment Schedule (WHO DAS II) in a primary care setting. *J Clin Epidemiol*. 2003;56(6):507-514. doi:10.1016/S0895-4356(03)00051-9
30. Hakkaart-van Roijen L, van Straten A, Donker M, Tiemens B. *Trimbos/IMTA Questionnaire for Costs Associated with Psychiatric Illness (TIC-P)*.; 2010.
31. Plaisier I, Beekman ATF, de Graaf R, Smit JH, van Dyck R, Penninx BWJH. Work functioning in persons with depressive and anxiety disorders: the role of specific psychopathological characteristics. *J Affect Disord*. 2010;125(1-3):198-206. doi:10.1016/j.jad.2010.01.072
32. Bond GR, Drake RE, Becker DR. Generalizability of the Individual Placement and Support (IPS) model of supported employment outside the US. *World Psychiatry*. 2012;11(1):32-39. doi:10.1016/j.wpsyc.2012.01.005
33. Finocchi C, Villani V, Casucci G, Villani V, Finocchi C. Therapeutic strategies in migraine patients with mood and anxiety disorders: Clinical evidence. *Neurol Sci Off J Ital Neurol Soc Ital Soc Clin Neurophysiol*. 2010;31 Suppl 1:S95-S98. doi:10.1007/s10072-010-0297-2
34. Linder J, Schüldt Ekholm K, Brodda Jansen G, Lundh G, Ekholm J. Long-term sick leavers with difficulty in resuming work: comparisons between psychiatric–somatic comorbidity and monodiagnosis. *Int J Rehabil Res*. 2009;32(1):20-35. doi:10.1097/MRR.0b013e328306351d
35. Cosci F, Fava GA. Staging of mental disorders: systematic review. *Psychother Psychosom*. 2013;82(1):20-34. doi:10.1159/000342243

36. Scott J, Leboyer M, Hickie I, et al. Clinical staging in psychiatry: a cross-cutting model of diagnosis with heuristic and practical value. *Br J Psychiatry*. 2013;202(4):243-245. doi:10.1192/bjp.bp.112.110858
37. Knight M, Stewart-Brown S, Fletcher L. Estimating health needs: the impact of a checklist of conditions and quality of life measurement on health information derived from community surveys. *J Public Health (Bangkok)*. 2001;23(3):179-186. doi:10.1093/pubmed/23.3.179
38. Kuh D, Shlomo Y Ben. *A Life Course Approach to Chronic Disease Epidemiology*. 2nd ed. New York: Oxford University Press; 2005.

Supplementary materials

Supplementary methods

We performed a number of sensitivity analyses in order to check for possible bias. First, the use of antidepressants could be a confounder for presence of CSD in ADD patients. We performed logistic regression analysis for presence of CSD categories in patients with ADDs, while adjusting for antidepressant use. In these analyses, antidepressant use was determined by inspecting medication containers and by referencing Anatomical Therapeutic Chemical (ATC) classifications. Antidepressant use was deemed present in those taking selective serotonin reuptake inhibitors (ATC code N06AB), tricyclic antidepressants (N06AA) or other antidepressants (N06AF/N06AX) at least 50% of the time.

In addition, a self-report bias could exist in which patients with ADDs report more CSDs than controls. In order to examine this possible self-report bias, we repeated our analyses with a more stringent operationalization of CSDs by checking whether current medication containers were in accordance with the CSD category respondents reported, as done previously [26]. In this more stringent operationalization, in addition to the self-report of certain CSDs, appropriate medication use was necessary for presence of a certain CSD category. Furthermore, frequency of medication use had to be at least 50% of the time, unless stated otherwise. See Supplementary Table I for an overview of Anatomical Therapeutic Chemical (ATC) classifications used in assessing medication-controlled CSDs.

In this way we identified 377 respondents who reported presence of CSDs but who did not use appropriate medications. We excluded these respondents from post hoc analyses in order to increase contrast between respondents with and without CSDs to be able to determine whether a more stringent operationalization of CSDs altered our findings. We repeated logistic regression analysis to assess differences in odds for CSD categories in patients with ADDs. Moreover, we repeated multivariate linear regression analysis to examine whether impact of CSDs, ADDs, and PM-comorbidity on disability differed after these adjustments for possible self-report bias. Finally, we repeated multivariate nominal logistic regression analyses to assess the impact of medication-controlled CSDs, ADDs, and their interaction on absenteeism and presenteeism.

Supplementary Table I. Medication controlled CSD categories.

CSD category		ATC code
Respiratory	Asthma, chronic bronchitis, pulmonary emphysema, other lung diseases	R01 (nasal preparations), R03 (Medication for obstructive airway diseases), R05 (Cough and cold preparations), R06 (Antihistaminics for systemic use), R07 (Other respiratory system products), H02 (Corticosteroids for systemic use)
Cardio-metabolic	Hypertension, angina pectoris, history of cardiac disease, stroke, diabetes, vascular abnormalities	C02 (antihypertensives), C03 (diuretics), C07 (beta blocking agents), C08 (calcium channel blockers), C09 (agent a/o renin-angiotensin system), C10 (Lipid-modifying agents), C01DA* (Nitrate vasodilators), B01 (Anti-coagulant / Anti-platelet agents), N02BA15 (Anti-coagulant / Anti-platelet agents), N02BA01 (Anti-coagulant / Anti-platelet agents), A10 (Medication used in diabetes),
Musculoskeletal	Osteoarthritis, rheumatoid arthritis, systemic lupus erythematodes, fibromyalgia, RSI, congenital skeletal deformation	M01 (Anti-inflammatory and anti-rheumatic products), N02A (Opioids), N02B (Other analgesics and antipyretics), H02 (Corticosteroids for systemic use), L04** (Immunosuppressants)
Gastrointestinal	Ulcer, irritable bowel syndrome, Crohn's disease, colitis ulcerosa, diverticulitis, liver cirrhosis, hepatitis, constipation, oesophageal disease, gastric sphincter dysfunction, other gastrointestinal disease	A02 (Medication for acid related disorders), A03 (Medication for functional gastrointestinal disorders), A04 (Antiemetics and antinauseants), A05 (Bile and liver therapy), A06 (laxatives), A07 (Antidiarrheals-intestinal anti-inflammatory/ anti-infective agents), H02 (Corticosteroids for systemic use), L04** (Immunosuppressants)
Neurological	Migraine, epilepsy, multiple sclerosis, peripheral neuropathy, hernia	N02 (Analgesics), M01A (Non-steroidal anti-inflammatory and anti-rheumatics products), M01B (Anti-inflammatory and anti-rheumatic agents in combination), N03 (Antiepileptics), H02 (Corticosteroids for systemic use), L03AB02 (Interferons)
Endocrine	Thyroid dysfunction	H03 (Medication used in thyroid dysfunction)
Cancer	Throat, thyroid, lymphoid, lung, oesophagus, bowel, stomach, liver, uterus, cervix, ovary, bladder, testicle, prostate, skin, brain, blood	L01/L02/L03/L04 (Medication used in cancer treatment), N02A (Opioids), N02B (Other analgesics and antipyretics), M01A (Non-steroidal anti-inflammatory and anti-rheumatics products), M01B (Anti-inflammatory and anti-rheumatic agents in combination)

* when taken at least 'if necessary'
** when taken <50% of time

Finally, we repeated multivariate linear regression analyses on total disability scores using different categorical variables for CSDs (no CSD, 1 CSD, 2+ CSDs) in interaction with ADDs; and for ADDs (no ADD, pure ADDs, comorbid ADDs) in interaction with CSDs. In this analysis, pure ADDs were defined as presence of either anxiety disorder or depressive disorder and comorbid ADDs were defined as presence of both anxiety and depressive disorders. We assessed separate main effects and interaction effects. We

formulated the following interaction terms: "1 CSD*ADD", "2+ CSD*ADD" (Table V, Model 2), and "pure ADD*CSD", "comorbid ADDs*CSD" (Table VI, Model 2).

Supplementary results and discussion

Associations between any CSD and ADDs remained the same after adjusting for antidepressant use (OR=1.32, 95%-CI: 1.06-1.64). Antidepressant use was more widespread in patients with ADDs, but was not accountable for differences in presence of CSDs.

Supplementary Table II shows associations of ADD patients with presence of medication-controlled CSDs. Odds for presence of any medication-controlled CSD are comparable to odds for any (uncontrolled) CSD (OR=1.31, 95%CI: 1.01-1.69). Odds for CSD categories are generally comparable, although odds were raised in those with ADD of having medication-controlled musculoskeletal disease (OR=1.81, 95%CI: 1.00-3.26) and neurological disease (OR=3.24, 95%CI:1.13-9.30). Still, in our original analyses, odds for presence of musculoskeletal and neurological diseases were raised in those with ADDs (OR=1.33, 95%CI: 0.95-1.85 for musculoskeletal disease and OR=1.78, 95%CI: 0.98-3.22 for neurological diseases), with p-values just above significance (p=.101 and p=.057, respectively). We therefore deemed these small differences as inconsequential with regard to our hypotheses.

Supplementary Table III shows that disability was raised more in those with ADDs (β=19.8) than those with medication-controlled CSDs (β=4.22), and interaction effects remained the same (β interaction=4.49) when using medication-controlled CSDs. Likewise, results for disability domains were comparable when using medication-controlled CSDs.

Supplementary Table IV shows that work impairment outcomes when using medication-controlled CSDs are generally comparable to work impairment outcomes when using non medication-controlled CSDs. Most notably, odds for all forms of work impairment are raised in those with PM-comorbidity (OR short absenteeism=2.96, 95%CI: 1.84-4.76; OR extended absenteeism=7.13, 95%CI:4.18-12.2; OR minor presenteeism=2.04, 95%CI: 1.25-3.31; OR extended presenteeism=16.8, 95%CI: 8.81-32.0) when compared to those with ADDs (without medication-controlled CSD) and those with medication-controlled CSDs (without ADDs). Therefore, applying a more stringent definition for presence of chronic somatic diseases did not alter our findings.

Supplementary Table II. Odds of having presence of Chronic Somatic Diseases (CSDs) (medication controlled) for Anxiety and/or Depressive Disorder (ADD) patients in comparison with a reference group without lifetime prevalence of ADDs.

		No lifetime ADD*	**Current ADD***		
		number/ *n*‡ (%)	number/ *n*‡ (%)	OR (95% CI)	p
Any chronic somatic disease n=595	Any of the listed below, medication controlled	142/ 549 (25.9%)	453/ 1445 (31.3%)	**1.31 (1.01-1.69)**	**.039**
Respiratory n=116	Asthma, chronic bronchitis, pulmonary emphysema, other lung diseases	23/ 614 (3.7%)	93/ 1657 (5.6%)	1.42 (0.88-2.28)	.151
Cardio-metabolic n=353	Hypertension, angina pectoris, history of cardiac disease, stroke, diabetes, vascular abnormalities	98/ 627 (15.6%)	255/ 1711 (14.9%)	1.05 (0.78-1.41)	.760
Musculoskeletal n=85	Osteoarthritis, rheumatoid arthritis, systemic lupus erythematodes, fibromyalgia, RSI, congenital skeletal deformation	15/ 596 (2.5%)	70/ 1619 (4.3%)	**1.81 (1.00-3.26)**	**.049**
Gastrointestinal n=109	Ulcer, irritable bowel syndrome, Crohn's disease, colitis ulcerosa, diverticulitis, liver cirrhosis, hepatitis, constipation, oesophageal disease, gastric sphincter dysfunction, other gastrointestinal disease	13/ 621 (2.1%)	96/ 1620 (5.9%)	**3.01 (1.65-5.49)**	**<.001**
Neurological n=39	Migraine, epilepsy, multiple sclerosis, peripheral neuropathy, hernia	4/ 623 (0.6%)	35/ 1708 (2.0%)	**3.24 (1.13-9.30)**	**.029**
Endocrine n=49	Thyroid dysfunction	14/ 630 (2.2%)	35/ 1718 (2.0%)	0.90 (0.47-1.73)	.755
Cancer n=28	Throat, thyroid, lymphoid, lung, oesophagus, bowel, stomach, liver, uterus, cervix, ovary, bladder, testicle, prostate, skin, brain, blood	5/ 599 (0.8%)	23/ 1642 (1.4%)	1.62 (0.59-4.43)	.349

All OR's were adjusted for gender, age and education.
* Anxiety and/or Depressive Disorders were defined as presence of either an Anxiety Disorder (Social Phobia, Generalized Anxiety Disorder, Panic Disorder with or without Agoraphobia) or a Depressive Disorder (Dysthymic of Major Depressive Disorder)
‡ *n* in psychiatric groups vary as we excluded those with self reported chronic somatic disease but without being prescribed adequate medication from each somatic disease category.

Supplementary Table V shows that the interaction effect of CSDs and ADDs on total disability does not differ when dividing CSDs into different categories (0 CSDs vs 1 CSD vs 2+ CSDs), as both 1 CSD and 2+CSDs interaction terms with ADD yielded significant positive interaction coefficients. Thus, the interaction effect between CSDs and ADDs is not driven by those with few or multiple CSDs, but can be generalized.

Supplementary Table III. Adjusted regression coefficients for the association between medication controlled Chronic Somatic Diseases (CSDs) and Anxiety and/or Depressive Disorders (ADDs) on total disability and disability domains (n=1,994).

Level of disability (WHO-DAS II)¶	Total score (32 items)	Cognition	Mobility	Self-care	Interpersonal interactions	Household activities	Participation in society
	β	β	β	β	β	β	β
Model 1, main effects							
Constant	**12.3****	**18.8****	**9.53****	**13.3****	**11.0****	**6.62**	**13.7****
Any CSD, medication controlled	**4.22****	**1.49**	**7.56****	**4.28****	**2.37***	**6.00****	**4.04****
ADDs***	**19.8****	**22.6****	**11.8****	**10.5****	**24.9****	**25.2****	**24.1****
Model 2, main effects and interaction effect							
Constant	**12.9****	**18.9****	**10.9****	**13.9****	**11.5****	**7.75***	**13.9****
Any CSD	0.79	0.80	0.20	0.90	-0.39	0.00	2.98
ADDs***	**18.6****	**22.4****	**9.17****	**9.31****	**23.9****	**23.0****	**23.7****
Interaction term CSD*ADD‡	**4.49****	0.91	**9.63****	**4.42***	3.61	**7.85****	1.39

All regression models were controlled for socio-demographics (gender, education, and age), β are unstandardized regression coefficients.
¶ WHO-DAS total and domain scores were standardized to values 0-100, a higher score indicates higher severity of disability.
*: p<.05, **: p<.01.
*** Anxiety and/or Depressive Disorders were defined as presence of either an Anxiety Disorder (Social Phobia, Generalized Anxiety Disorder, Panic Disorder with or without Agoraphobia) or a Depressive Disorder (Dysthymic or Major Depressive Disorder)
‡ in this model, interactions are tested on an additive scale, a significant positive interaction effect (for instance: β =4.49) implies that the effect of comorbid CSD and ADD on disability score is larger than the sum of separate effects for CSD and ADD.

In a similar sensitivity analysis, supplementary Table VI shows that the interaction terms for two subgroups of ADD severity (no ADDs vs pure ADD (either anxiety or depressive disorder present) vs comorbid ADDs (anxiety and depressive disorders present) are both positive and significant. So, the effect between CSDs and ADDs on total disability is not driven by those with comorbid ADDs (those with both anxiety and depressive disorders), or by those with pure disorders.

Supplementary Table IV. Multinomial regression coefficients for impact of medication controlled Chronic Somatic Diseases (CSDs) and current Anxiety and/or Depressive Disorders (ADDs) on absenteeism in employed respondents (n=1,230).

		Absenteeism§				**Presenteeism‡**			
		Short absenteeism (<2 weeks) vs. no absenteeism (ref)		**Extended absenteeism (>2 weeks) vs. no absenteeism (ref)**		**Minor presenteeism vs. no presenteeism (ref)**		**Extended presenteeism vs. no presenteeism (ref)**	
	n	*OR*	*(95% CI)*	*OR*	*(95% CI)*	*OR*	*(95% CI)*	*OR*	*(95% CI)*
Model 1, main effects									
Any CSD	305	1.09	(0.76-1.57)	1.12	(0.78-1.59)	1.17	(0.81-1.70)	**1.63**	**(1.15-2.32)****
ADDs***	864	**2.74**	**(2.02-3.72)****	**6.42**	**(4.37-9.42)****	**1.77**	**(1.31-2.39)****	**8.19**	**(5.32-12.6)****
Model 2, main effects									
Controls *No CSDs, no ADDs*	281	1.00	(ref)	1.00	(ref)	1.00	(ref)	1.00	(ref)
Purely physical *CSD present, no ADDs*	85	1.24	(0.66-2.34)	1.19	(0.53-2.67)	1.22	(0.65-2.32)	**3.89**	**(1.69-8.97)****
Purely mental *no CSDs, ADD present*	644	**2.85**	**(2.02-4.03)****	**6.57**	**(4.20-10.3)****	**1.81**	**(1.29-2.54)****	**11.7**	**(6.59-20.9)****
Physical mental comorbidity *CSD and ADD present*	220	**2.96**	**(1.84-4.76)****	**7.13**	**(4.18-12.2)****	**2.04**	**(1.25-3.31)****	**16.8**	**(8.81-32.0)****

These regression models were controlled for socio-demographics (gender, education, age).
§ absenteeism is represented by number of weeks absent from work during the last six months.
‡ presenteeism is represented by number of workweeks in which quality of work was reduced due to health issues, multiplied by a self-reported proportional score for severity of work quality reduction.
*****: p<.05, ******: p<.01.
*** Anxiety and/or Depressive Disorders were defined as presence of either an Anxiety Disorder (Social Phobia, Generalized Anxiety Disorder, Panic Disorder with or without Agoraphobia) or a Depressive Disorder (Dysthymic or Major Depressive Disorder)

Supplementary Table V. Sensitivity analysis to assess whether the interaction effects between CSDs and ADDs persist with increasing number of CSDs (*n*=2,371).

Level of disability (WHO-DAS II)[a]		
	n	*β*
Model 1, main effects		
constant		**12.8****
1 CSD (reference 0 CSDs)	657	**2.74****
2+ CSDs (reference 0 CSDs)	312	**6.69****
ADDs[b]	1,737	**20.0****
Model 2, main effects and interaction effect		
constant		**13.5****
1 CSD (reference 0 CSDs)	657	0.36
2+ CSDs (reference 0 CSDs)	312	2.92
ADDs[b]	1,737	**18.6****
Interaction term 1 CSD*ADD[c]	485	**4.65***
Interaction term 2+CSD*ADD[c]	258	**3.21***

These regression models were controlled for socio-demographics (gender, education, and age), β are unstandardized regression coefficients.

[a] WHO-DAS total and domain scores were standardized to values 0-100, a higher score indicates higher severity of disability.

***: p<.05, ****: p<.01.

[b] Anxiety and/or Depressive Disorders were defined as presence of either an Anxiety Disorder (Social Phobia, Generalized Anxiety Disorder, Panic Disorder with or without Agoraphobia) or a Depressive Disorder (Dysthymic or Major Depressive Disorder)

[c] in this model, interactions are tested on an additive scale, a non-significant positive interaction effect (for instance: β interaction =4.65) implies that the effect on disability score of comorbid ADDs in those with 1 CSD is larger than the sum of separate effects for CSD and ADD (i.e. synergistic effect modification).

Supplementary Table VI. Sensitivity analysis to assess whether the interaction effects between CSDs and ADDs persist with increasing number of ADDs (n=2,371).

Level of disability (WHO-DAS II)[a]		
	n	*β*
Model 1, main effects		
constant		**10.1****
CSDs	969	**3.54****
1 ADD[b] (reference 0 ADDs)	947	**14.9****
Comorbid ADDs (anxiety and depression, reference 0 ADDs)	790	**26.9****
Model 2, main effects and interaction effect		
constant		**10.8****
CSDs	969	0.79
1 ADD[b] (reference 0 ADDs)	947	**13.7****
Comorbid ADDs (anxiety and depression, reference 0 ADDs)	790	**25.2****
CSD*1 ADD	387	**3.17***
CSD*comorbid ADDs	356	**4.15****

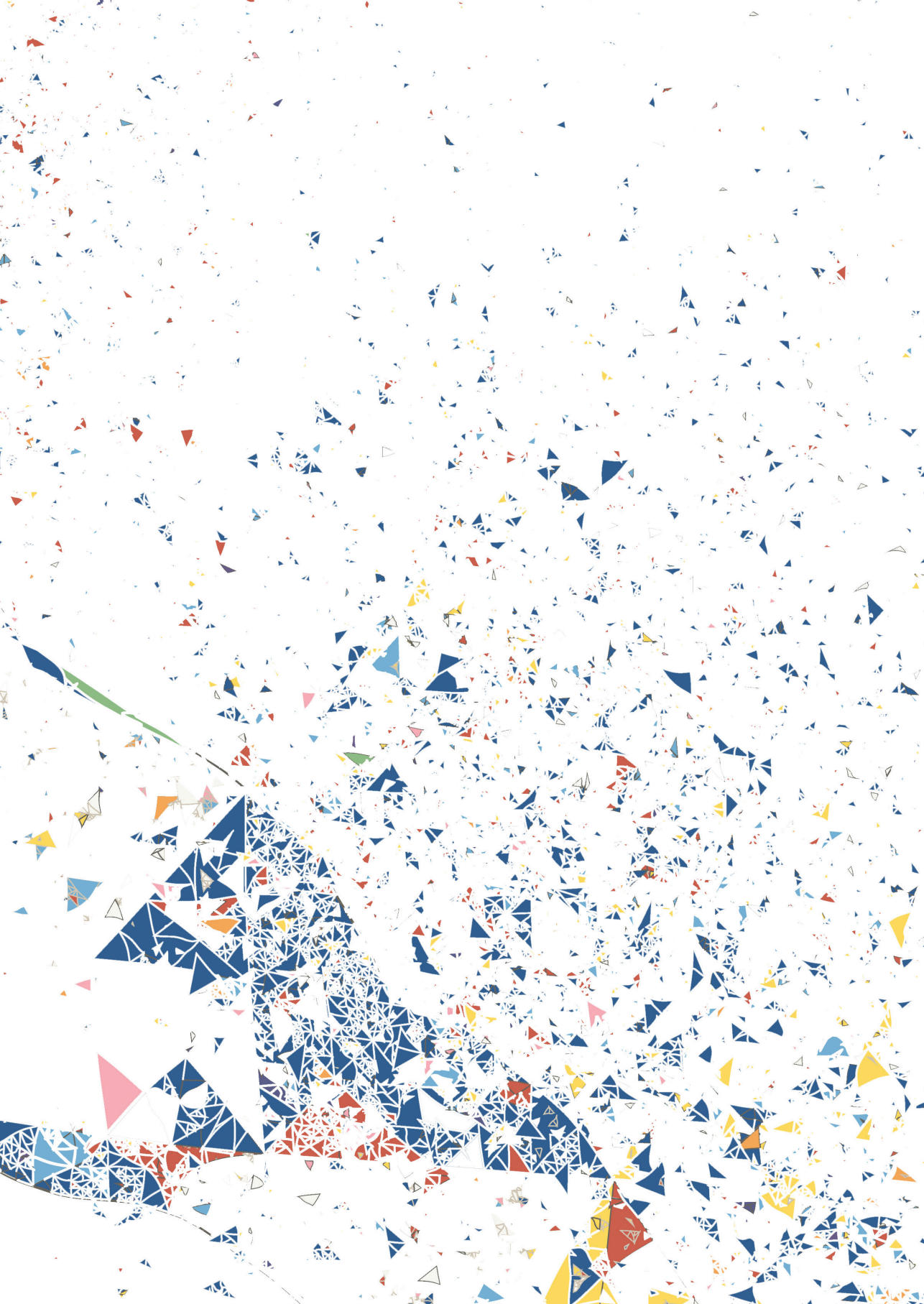
These regression models were controlled for socio-demographics (gender, education, and age) , β are unstandardized regression coefficients.

[a] WHO-DAS total and domain scores were standardized to values 0-100, a higher score indicates higher severity of disability.

*: p<.05, **: p<.01.

[b] Anxiety and/or Depressive Disorders were defined as presence of either an Anxiety Disorder (Social Phobia, Generalized Anxiety Disorder, Panic Disorder with or without Agoraphobia) or a Depressive Disorder (Dysthymic or Major Depressive Disorder)

[c] in this model, interactions are tested on an additive scale, a significant positive interaction effect (for instance: β *interaction* =4.15) implies that the effect of comorbid CSD and ADD on disability score is larger than the sum of separate effects for CSD and ADD (i.e. synergistic effect modification).



Chapter 4.

Aligning the many definitions of treatment resistance in Anxiety Disorders: a systematic review

Depression and Anxiety (2019)

Wicher A. Bokma
Guido A.A.M. Wetzer
Jurriaan B. Gehrels
Brenda W.J.H. Penninx
Neeltje M. Batelaan
Anton L.J.M. van Balkom

Abstract

Anxiety Disorders often show a chronic course, even when treated with one of the various effective treatments available. Lack of treatment effect could be due to Treatment Resistance (TR). Consensus on a definition for TR Anxiety Disorders (TR-AD) is highly needed as currently many different operationalizations are in use. Therefore, generalizability in current TR-AD research is suboptimal, hampering improvement of clinical care. The objective of this review is to evaluate the currently used definitions of TR-AD by performing a systematic review of available literature. Out of a total of n=13,042, 62 studies that operationalized TR-AD were included. The current review confirms a lack of consensus on TR-AD criteria. In 62.9% of the definitions, TR was deemed present after the first treatment failure. Most studies (93.0%) required pharmacological treatment failures, whereas few (29.0%) required psychological treatment failures. However, criteria for what constitutes "treatment failure" were not provided in the majority of studies (58.1%). Definitions for minimal treatment duration ranged from at least four weeks to at least six months. Almost half of the TR-AD definitions (46.8%) required elevated anxiety severity levels in TR-AD. After synthesis of the results, the consensus definition considers TR-AD present after both at least one first-line pharmacological and one psychological treatment failure, provided for an adequate duration (at least eight weeks) with anxiety severity remaining above a specified threshold. This definition could contribute to improving course prediction and identifying more targeted treatment options for the highly burdened subgroup of TR-AD patients.

Introduction

Up till now, a widely used definition for treatment resistance in Anxiety Disorders does not exist.[1-5] This is surprising because it is well known that a substantial proportion of adults with Anxiety Disorders experience suboptimal treatment results after evidence-based treatments.[6-8] Selective Serotonin Reuptake Inhibitors (SSRIs) and Serotonin Norepinephrine Reuptake Inhibitors (SNRIs) are widely regarded the first-line pharmacological treatments for Anxiety Disorders.[9-11] Cognitive Behavioural Therapy (CBT) is the psychological first-line treatment option for Anxiety Disorders.[12] First-line treatments show a moderate effect size in meta-analytic comparisons with placebo.[13,14] After first-line treatment up till 30-60% of patients have substantial and impairing remaining symptoms.[7,9,15]

A wide variety of terms are in use for the phenomenon of suboptimal treatment results in anxiety disorders: "refractory anxiety", "treatment resistance", "medication resistance", "treatment refractory cases", "remaining symptomatic" and "persistent symptoms".[16-21] In other psychiatric disorders the term "treatment resistance" (TR) is preferred to describe a subgroup of patients who have a prior history of unfavorable treatment effects,[22-25] which also implies having less favorable future treatment effects.[22-25] The varying terminology reflects the absence of consensus regarding the *criteria* for TR.[1,5,26-28] This lack of consensus on criteria for TR-AD was first recognized in 2004; however, fourteen years later still no consensus exists.[1-5,29]

Most authors define Treatment Resistant Anxiety Disorders (TR-AD) as the persistence of anxiety symptoms, or as the absence of response, recovery or remission of the disorder after some form of active treatment.[1,30-34] These active treatments should represent evidence-based treatment regimes, provided at an adequate dosage and for an adequate duration.[3,35] However, the absence of anxiety symptoms does not always indicate full disorder remission.[2,36] A substantial amount of residual disease burden may be present in persisting behavioural changes such as avoidance, or in altered cognitive functioning, for instance in excessive rumination. Additional emphasis on functional recovery is therefore advocated by a number of authors when assessing TR-AD.[2,36] No systematic review into the definition for TR-AD is yet performed.

The aim of this study is to summarize and discuss the different criteria used for TR-AD. In order to do this, we will perform a systematic literature review. Second, by summarizing and comparing the different criteria used for TR in anxiety disorders, we aim to propose a consensus definition for TR-AD.

Methods

The methods for this systematic review were specified in advance in a study protocol which was documented in the PROSPERO database (reference number CRD42017055864). The current paper was drafted in accordance with the PRISMA guidelines for reporting on systematic reviews.[37]

Literature search

A systematic search across MEDLINE, PubMed (non-MEDLINE), EMBASE, PsycINFO, and Web of Science for available literature until April 2018 was performed. In order to derive all articles that might include a definition for TR in anxiety disorders we searched for Anxiety Disorders (according to DSM-5)[38] in combination with various free-text synonyms for 'treatment resistance' (see Panel 1 for the full search query).

All publication types in English were included with the exception of conference summaries, editorials, columns, book reviews and manifestos as these were unlikely to include a full description of a TR-AD definition. Studies were selected when they included adults or elderly persons with anxiety disorders (Panic Disorder, with or without Agoraphobia, PD(A), Social Anxiety Disorder (SAD), Generalized Anxiety Disorder (GAD), Specific Phobia (SP), Selective Mutism, and Separation Anxiety). No restrictions in presence of comorbidity were used. Exclusion criteria included studies with an average study population below 21 years, and studies reporting primarily on Post-Traumatic Stress Disorder (PTSD) or Obsessive-Compulsive Disorder (OCD), because these are no longer classified as Anxiety Disorders.

Panel 1.

Overview of search terms used in this systematic review (formatted for MEDLINE)

((“Anxiety Disorders”[Mesh:NoExp] OR “Agoraphobia”[Mesh] OR “Anxiety, Separation”[Mesh] OR “Neurocirculatory Asthenia”[Mesh] OR “Neurotic Disorders”[Mesh] OR “Panic Disorder”[Mesh] OR “Phobic Disorders”[Mesh] OR anxiety disorder* [tiab] OR generalized anxiety disorder* [tiab] OR generalised anxiety disorder* [tiab] OR anxiety state* [tiab] OR agoraphobi* [tiab] OR panic* [tiab] OR phobi* [tiab] OR selective mutis* [tiab]))

AND

(“Retreatment” [Mesh] OR “Drug Resistance” [Mesh:NoExp] OR “Drug tolerance” [Mesh] OR treatment resistan* [tiab] OR refractor* [tiab] OR poor respon* [tiab] OR partial respon* [tiab] OR non-respon* [tiab] OR nonrespon* [tiab] OR loss of respons* [tiab] OR medication resistan* [tiab] OR drug resistan* [tiab] OR tachyphyl* [tiab] OR resilien* [tiab] OR persistan* [tiab] OR immune [tiab] OR insusceptib* [tiab] OR irresponsive* [tiab] OR unreceptive* [tiab] OR resistive [tiab] OR unsuccessful treatment* [tiab] OR treatment failur* [tiab] OR failed treatment* [tiab] OR “Patient Dropouts”[Mesh] OR patient dropout* [tiab] OR treatment dropout* [tiab] OR “Patient Compliance”[Mesh] OR non-complian* [tiab] OR noncomplian* [tiab] OR non-adheren* [tiab] OR nonadheren* [tiab] OR remaining symptom* [tiab] OR pseudo-resistan* [tiab] OR dropping out [tiab] OR augmentation [tiab] OR inadequate respon* [tiab] OR intractab* [tiab] OR partially respon* [tiab] OR resistant patient* [tiab] OR remain symptom* [tiab] OR remaining symptom* [tiab] OR non-remitting [tiab] OR nonremitting [tiab] OR partial improvement* [tiab] OR incomplete respon* [tiab] OR residual symptom* [tiab] OR anxiolytic toleran* [tiab])

Eligibility assessment

Eligibility assessment on title and abstract was performed independently by two reviewers (WB, GW, JG) by using the Cochrane-supported review program Covidence (www.covidence.org). Disagreements were resolved

by consensus after discussion. A flow chart for inclusion of eligible studies according to PRISMA guidelines is provided in figure 1. Full-text screening was performed independently by two reviewers (WB and JG). During the full-text screening phase, articles were excluded if a full-text version could not be retrieved or if any of the exclusion criteria were present. Studies were included if their definition for TR-AD could be implicitly deduced from inclusion criteria used in a study. Reviews, meta analyses and book chapters were included if they provided their own definition for TR-AD but were excluded if they repeated other studies' definitions without providing rationale for choosing this definition over others. As the vast majority of studies used TR and "refractory" interchangeably we chose to regard them as synonyms and will refer to these phenomena as TR-AD.

Data extraction

From trials we extracted data on study characteristics: number of subjects, population of interest, intervention, comparator condition, follow-up period, primary outcomes and results; from reviews we extracted data on study design and population of interest. With regard to the definitions for TR-AD, we extracted data on nine predefined putative criteria for the definition, based on criteria used in the Maudsley Staging Method for treatment resistant depressive disorders.[23] Additionally, we extracted one TR-AD criterion (treatment response), that was not predefined in our study protocol. The ten criteria were: minimal number of failed treatments, failed psychotherapy trials, failed pharmacological trials, failed other biological treatments, minimal length of treatment, treatment response criterion (i.e. which post-treatment change constitutes response/ failure), minimal duration of anxiety disorder, severity of symptoms, presence of functional impairment, and presence of comorbidity. We evaluated which of these ten criteria were present in TR-AD definitions across included studies (yes/no). Specific values for each criterion were extracted as well.

Quality of definitions

We assessed the definition quality in each included study. As there are no formal risk of bias tools available for the purpose of our study, and as we are not interested in potential sources of study outcome bias we assessed definition quality in two ways; first, by counting the total number of TR-AD criteria included in each study's definition, second, by determining the degrees of precision with which the definition for TR-AD is presented in each paper. The total number of TR-AD criteria was a count variable counting presence of all ten dichotomized TR-AD criteria. Degrees of precision was

categorized into 'high', 'medium' and 'low'. Precision was considered 'high' if a study provided an explicit definition for TR-AD, for example in this study by De Salas-Cansado et al. (2013):

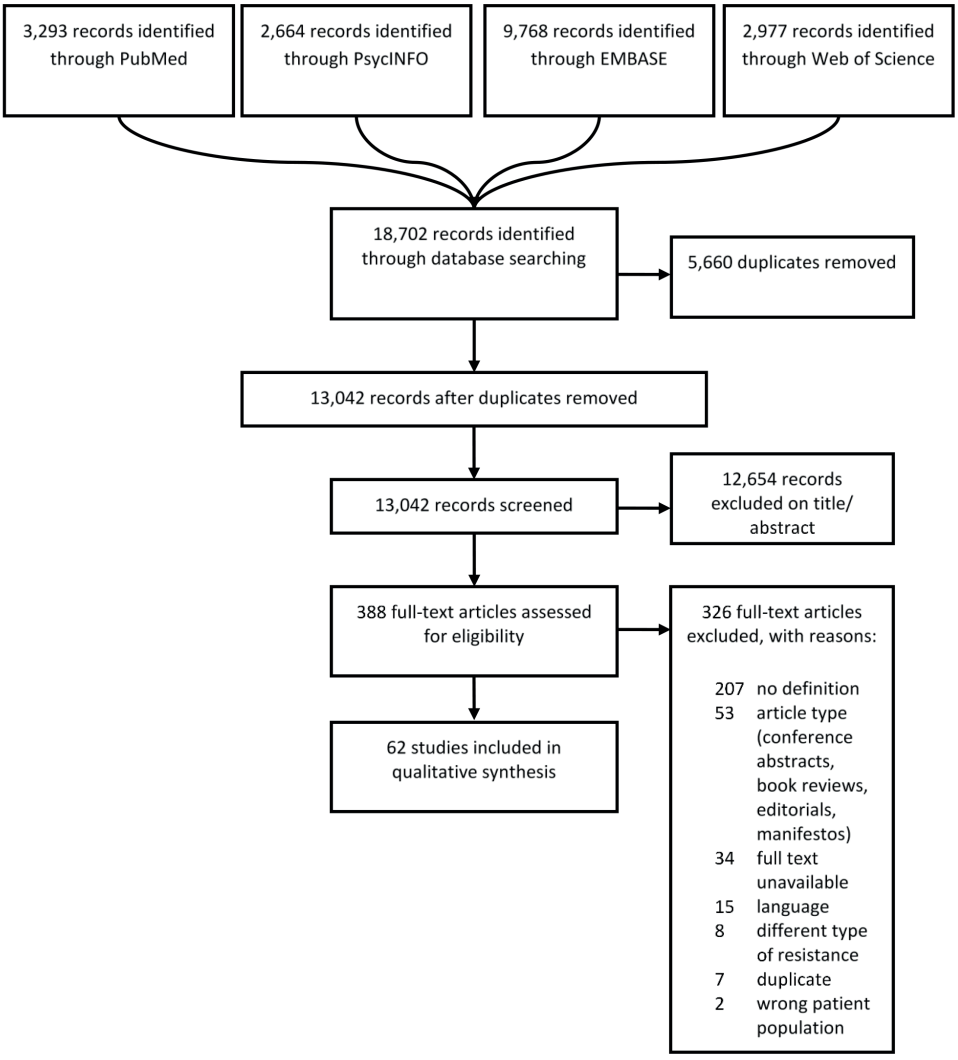


Figure 1. PRISMA flow chart for study inclusion.

> *Refractory was defined as subjects with persistent symptoms/ suboptimal response, a Hamilton-anxiety (HAM-A) scale score ≥16 and a Clinic Global Impression (CGI) score ≥3 at baseline, after a standard dose regimen of any anti-anxiety drug, alone or in combination, for at least 6 months, given prior to the baseline study visit. (p987)*[39]

The degree of precision was deemed 'medium' if the criteria were only implicitly attributable to the concept of TR-AD, or if multiple terms were used interchangeably, for instance in a study by Lohoff, Etemad, Mandos, Gallop, & Rickels (2010) in patients with "refractory GAD":

> *Subjects also had to have treatment failure of at least 1 adequate trial of an SSRI, an SNRI, a BZ, or a combination of these agents. Patients who were on an SSRI, an SNRI, a BZ, or a combination of these agents before enrollment had to be on a stable dose for 4 weeks. Inclusion further required a total score of 16 or higher on the Hamilton Anxiety Scale (HAM-A) and a score of 4 or greater on the Clinical Global Impression Severity of Illness Scale (CGI-S) (p186).*[40]

Finally, if the study only provided a description of the concept of TR-AD, without operationalizing it in specific criteria, the degree of precision was deemed 'low', for instance:

> *"failure of an adequate clinical trial of medication" (Stein, 2004).*[30]

Data synthesis

In order to synthesize the results of the systematic review into a new operationalization for TR-AD, frequencies for presence of each individual TR-AD criterion were assessed. The most frequently used values for each individual criterion were considered the most appropriate operationalization for that criterion and were chosen for the consensus definition. However, if an unspecified category for a certain criterion (for example "unspecified type of pharmacological treatment") was the most frequently used value, we did not consider this category for the new definition if a more specified value was available. Additionally, criteria that were included only in a small minority (<10%) of the studies were not used for the new definition, as they were then judged to be lacking a convincing empirical basis.

Statistical analyses

To test associations between total number of criteria provided in definitions, degrees of precision and publication year, we performed Kruskall-Wallis tests for differences in mean rank. We hypothesized that higher definition quality studies (i.e. more total criteria or a higher degree of precision) would be the most recent studies. Fisher's Exact tests were performed to investigate whether the two definition quality variables were associated with different frequencies for values of each TR-AD criterion. For instance: did high definition quality studies more often require a higher number of failed treatments or more often mention a SSRI/SNRI failure as requisite for TR-AD compared with lower definition quality studies?

Results

Study selection

The electronic database search yielded 18,702 results. After deduplication 13,042 entries remained. During title and abstract screening, 12,654 studies were excluded. We assessed 388 full-text studies, of which 207 did not contain a definition, 53 were a wrong article type (conference abstracts and editorials), 34 could not be retrieved, 15 did not meet language requirements, 8 reported on a different type of "resistance" (e.g. "resistance" in the psychodynamic paradigm), 7 were previously unrecognized duplicates and 2 reported on a different patient population. This resulted in the final inclusion of sixty-two studies (for a flow chart see Figure 1).

Characteristics of included studies

Included studies were published between 1986 and 2018. They consisted of eight narrative reviews,[2,4,9,34,36,41–43] five systematic reviews,[1,16,44–46] of which three also performed meta-analyses,[1,16,46] seven treatment guidelines/algorithms,[30,47–52] three book chapters,[53–55] twenty-one open-label trials,[26,56–75] eight RCTs,[40,76–82] four retrospective cohorts,[83–86] one prospective cohort,[87] three case series,[88–90] one cost-effectiveness analysis,[39] and one trial protocol.[32] Thirty-three studies pertained to PD, thirty-four to GAD, twenty-one to SAD, two to SP, and five to Anxiety Disorders in general. For a summary of study characteristics see Table 1. For full details, see eTable 1 (trials, cohort studies and meta-analyses) and eTable 2 (reviews, treatment guidelines and book chapters).

Table 1. Study characteristics for included studies.

Study characteristics	n	%
Publication type		
Book chapter	3	4.8
Case series	3	4.8
Cost-effectiveness analysis	1	1.6
Narrative review	8	12.9
Open-label trial	21	33.9
Prospective cohort study	1	1.6
Randomized controlled trial	8	12.9
Retrospective cohort study	4	6.5
Systematic review	2	3.2
Systematic review + meta-analysis	3	4.8
Trial protocol	1	1.6
Treatment guidelines/ algorithms	7	11.3
Population of interest[1]		
Anxiety disorders (in general)	5	8.1
Generalized anxiety disorder	34	54.8
Panic disorder	33	53.2
Social anxiety disorder	21	33.9
Specific phobia	5	8.1
Type of intervention used (if any)		
Adjunctive psychotherapy	8	12.9
Any therapy	1	1.6
Any adjunctive therapy	1	1.6
Combination treatment: pharmacological and psychological	4	6.5
Either pharmacologic monotherapy or pharmacologic augmentation therapy	1	1.6
Pharmacologic augmentation or combination treatment	17	27.4
Pharmacologic monotherapy	7	11.3
Nervus vagus stimulation	1	1.6
Self-management	1	1.6
Degree of precision of included definitions		
High	13	21.0
Medium	44	71.0
Low	5	8.1

[1] Some studies described more than one population of interest.

Definition quality

The total number of criteria per study ranged from 1 to 6 (mean=3.58, SD=1.31). With respect to the assessment of the degree of precision for TR-AD definitions it appeared that 13 studies (21.0%) provided a high degree of precision, 44 (71.0%) a medium degree, and 5 (8.1%) a low degree of precision.

There was a significant association between total number of criteria and year of publication (X^2(df=5)=13.01,p=0.02): the studies with the highest number of criteria were, on average, the most recent. For degrees of precision no association with publication date existed (X^2(df=2)=2.13,p=0.34). Neither studies with a higher total number of criteria, nor studies with a higher degree of precision provided a different perspective on the ten TR-AD criteria. Since definition quality did not change operationalizations for TR-AD, all studies were used in the synthesis of results.

Main results

By applying a systematic review approach it became apparent that a large majority of studies on the topic of TR-AD (n=207) do not provide a definition for the phenomenon of TR-AD. Furthermore, the included studies (n=62) yielded many different definitions for the concept of TR-AD (see eTable 3 for all definitions). Trials often used the presence of one failed pharmacological treatment as an adequate definition for TR-AD. Other studies provided additional criteria. When the frequencies for each of the ten extracted TR-AD criteria were compared across included studies, some distinctive patterns arose (see Table 2).

The minimal number of required failed treatments, regardless of treatment type, was reasonably consistent across studies: 39 studies (62.9%) required one treatment failure for TR-AD, with other studies varying between two (n=3) and five (n=1) failed previous treatments. Failed psychotherapy trials were only included in 18 studies (29.0%). These studies all regarded CBT an appropriate treatment, with seven studies (11.3%) restricting TR-AD to CBT failure alone, whilst others (n=9, 14.5%) also regarded other psychological treatments appropriate.

Contrastingly, a large majority (n=58, 93.5%) required at least one failed pharmacotherapy trial for their definition for TR-AD. Of these, some studies (n=15, 24.2%) considered at least one failed SSRI/SNRI trial sufficient to be classified as TR-AD. A substantial number of studies did not specify type of

pharmacotherapeutic treatment failure required for TR-AD (n=27, 43.5%), for instance by referring to "first-line" or "standard" antianxiety treatments. A few used a varying number of treatment types in a stepped-care or staging algorithm (n=2, 3.2%) or considered other pharmacotherapeutic treatment failures adequate (n=14, 22.6%). See eTable 3 for detailed descriptions of type of pharmacotherapy. Whether failed trials were caused by a lack of effect, or a lack of tolerability was usually not reported. Other biological treatments were not included in TR-AD definitions.

Most studies (n=34, 54.8%) used a minimal treatment length criterion ranging from four weeks to six months, while the most often used adequate minimal treatment duration was eight weeks (n=15, 24.2%).

A substantial number of studies (n=26, 41.9%) gave a response criterion. The most commonly used cut-off values were a <50% posttreatment improvement on the Hamilton Anxiety Rating Scale (HAM-A) and a posttreatment Clinical Global Impression Improvement scale (CGI-I) score greater than two (i.e. "minimal improvement", at best). Severity of anxiety symptoms was often included in definitions (n=29, 46.8%), with cut-off scores commonly provided: a HAM-A score of above 16 (for any Anxiety Disorder), a Clinical Global Impression Severity Scale (CGI-S) score of four or higher (for any Anxiety Disorder), a total score above 3, or any item above 1 on the Panic Disorder Severity Scale (PDSS) for PD and a score above 60 on the Leibowitz Social Anxiety Scale (LSAS) for SAD. For GAD, no disorder-specific measurement instrument was reported in TR-AD definitions. Finally, minimal disease duration (n=2, 3.2%), presence of functional impairments (n=5, 8.1%) and presence of comorbidity (n=1, 1.6%) were sparsely included in definitions for TR-AD. See Table 2 for a summary per TR-AD criterion, and eTable 4 for a full overview of included TR-AD criteria per study.

Synthesis of results

In order to propose a consensus definition for TR-AD that reflects the current literature, we included the most prevalent values for all criteria that were provided consistently across studies into the new TR-AD definition. Failed SSRI/SNRI trials were most often considered as criterion for TR-AD. Studies typically referred to SSRI/SNRI trials as 'first-line' treatment. Therefore, failure of at least one first-line treatment (SSRI/SNRI) was included in the new definition. Although psychotherapeutic treatment failure was less often incorporated in TR-AD definitions, CBT was usually referred to as 'first-line' psychological intervention. Therefore, first-line psychological interventions

(CBT) failure was included in the new definition. Although the most often provided criterion for minimal number of treatments was one, by including both pharmacological and psychological treatment failures into the definition, the minimal number of failed treatments in the new definition rose to at least two. A minimal adequate treatment duration of eight weeks was included in the consensus definition. In studies that permitted psychotherapy failures as criterion for TR-AD (n=21), only five provided a minimal treatment duration criterion, ranging from 4 weeks to 20 sessions CBT (see eTable 4). Therefore, a minimal duration of 8 weeks was maintained for psychotherapy trials.

Table 2. Criteria included in definitions for treatment resistance anxiety disorders.

Treatment resistance definition criteria	n	%
Minimal number of failed treatments		
Not part of definition	9	14.5
Included in definition	53	85.5
Unspecified or varying number	*7*	*11.3*
1 failed treatment	*39*	*62.9*
2 failed treatments	*3*	*4.8*
3 or more failed treatments	*4*	*6.5*
Failed psychotherapy trials		
Not part of definition	44	71.0
Included in definition	18	29.0
Any	*9*	*14.5*
At least one failed CBT trial	*7*	*11.3*
Varying number (stepped-care or staging approach)	*2*	*3.2*
Failed pharmacological trials		
Not part of definition	4	6.5
Included in definition	58	93.5
Unspecified number or type of failed pharmacological treatment	27	43.5
At least one failed SSRI/SNRI trial	*15*	*24.2*
At least one failed other pharmacotherapeutic trial	*14*	*22.6*
Varying number or types (stepped-care or staging approach)	*2*	*3.2*
Other biological treatments		
Not part of definition	62	100
Included in definition	0	0

Table continues

Table 2. Continued

Treatment resistance definition criteria	n	%
Minimal length of treatment		
Not part of definition	28	45.2
Included in definition	34	54.8
>4 weeks	*8*	*12.9*
>6 weeks	*3*	*4.8*
>8 weeks1	*15*	*24.2*
>11 weeks or 20 sessions of CBT	*1*	*1.6*
>12 weeks	*3*	*4.8*
>4 months	*2*	*3.2*
>6 months	*2*	*3.2*
Treatment response criterion		
Not part of definition	36	58.1
Included in definition	26	41.9
Cut-off values for effective/ failed treatment provided[2]	*26*	*41.9*
Minimal duration of anxiety disorder		
Not part of definition	60	96.8
Included in definition	2	3.2
>1 year	*1*	*1.6*
>2 years	*1*	*1.6*
Severity of symptoms		
Not part of definition	33	53.2
Included in definition	29	46.8
Aspecific criterion (e.g. "severe")	*1*	*1.6*
Specific criterion (cut-off values) provided[3]	*28*	*45.2*
Functional impairment		
Not part of definition	57	91.9
Included in definition	5	8.1
Aspecific criterion (e.g. "marked impairments")	*4*	*6.5*
Specific criteria (cut-off values) provided[4]	*1*	*1.6*
Presence of comorbidity		
Not part of definition	61	98.4
Included in definition	1	1.6
Comorbidity as exclusion criterion for TR-AD	*1*	*1.6*

[1] including studies with minimal treatment duration of '2 months'.
[2] the most often used criteria were: ΔHAM-A<50% or CGI-I>2
[3] the most often used criteria for severe symptomatology were HAM-A<16 or CGI-S≥4 (for all Anxiety Disorders), PDSS>3 or any PDSS item>1 (for PD(A)), LSAS>60 (for SAD)
[4] one study used SDS >1 on each item as criterion for functional impairments.
Abbreviations: CBT= Cognitive Behavioral Therapy, SSRI=Selective Serotonin Reuptake Inhibitor, SNRI= Selective Norepinephrine Reuptake Inhibitor. HAM-A: Hamilton Anxiety Rating Scale, CGI-I= Clinical Global Impression Improvement Scale, CGI-S= Clinical Global Impression Severity Scale, PDSS= Panic Disorder Severity Scale, LSAS= Leibowitz Social Anxiety Scale, PD(A)= Panic Disorder (with or without agoraphobia), SAD= Social Anxiety Disorder, SDS= Sheehan Disability Scale

Absence of treatment response was included in the consensus definition, using the two most commonly provided cut-off values from studies included in this review. Other biological treatments, minimal duration of anxiety disorder, presence of functional impairments and comorbidity were only sporadically included in TR-AD definitions and therefore were not considered for the consensus TR-AD definition. See Panel 2 for the full description of this consensus TR-AD definition with most commonly used cut-off values for each criterion.

Panel 2.

Proposed operationalization for Treatment Resistant Anxiety Disorders (TR-AD).

TR-AD checklist			
Failed pharmacotherapeutic treatment			
☐	At least one first-line treatment (SSRI, SNRI)[1]		
☐	pre-to posttreatment difference in HAM-A <50% or posttreatment CGI-I >2		
☐	treatment period of at least 8 weeks		
Failed psychotherapeutic treatment			
☐	At least one first-line psychotherapeutic treatment (CBT)[2]		
☐	pre-to posttreatment difference in HAM-A <50% or posttreatment CGI-I >2		
☐	provided according to local protocols and for an adequate duration (at least>8 weeks)		
Current severity of anxiety symptoms			
☐	GAD	HAM-A >15	or CGI-S > 3
☐	PD	HAM-A >15 or PDSS >3, or any item >1	or CGI-S > 3
☐	SAD	HAM-A >15 or LSAS ≥ 60	or CGI-S > 3

TR-AD is present if all six treatment boxes can be checked in addition to at least one symptom severity box
[1] SSRIs and SNRIs are considered first-line pharmacotherapeutic treatment options as per 2018[19,21,49]
[2] CBT interventions are considered first-line psychotherapeutic treatment options as per 2018[17,49]
Abbreviations: SAD= Social Anxiety Disorder, PD= Panic Disorder, GAD= Generalized Anxiety Disorder, HAM-A= Hamilton Anxiety Rating Scale, PDSS= Panic Disorder Severity Scale, LSAS= Leibowitz Social Anxiety Scale, CGI-S= Clinical Global Impression Severity Scale, CGI-I= Clinical Global Impression Improvement Scale, SSRI= Selective Serotonin Reuptake Inhibitor, SNRI= Selective Serotonin and Norepinephrine Reuptake Inhibitor, CBT= Cognitive Behavioral Therapy.

Discussion

This paper aimed to systematically review different definitions and criteria for treatment resistant Anxiety Disorders (TR-AD) and showed that the majority of studies do not provide a definition for TR-AD and that consistency

and consensus across TR-AD definitions in included studies is lacking. The most frequently used definition for TR-AD simply consists of one failed first-line pharmacotherapy treatment. Both the lack of consensus in current TR-AD definitions and the unclear description of TR-AD in the most used definition make the current attempt of aligning definitions a necessity.

Out of ten putative criteria, we identified six criteria that are regularly integrated into the various different definitions for TR-AD: minimal number of treatment failures, presence (and type) of psychological treatment failure, presence (and type) of pharmacological treatment failure, minimal treatment duration (>8 weeks), specification of a response criterion (i.e. what constitutes a "failed treatment"), and minimal symptom severity. These criteria were integrated into a consensus definition. Four putative criteria were dismissed: "minimal duration of disorder", "other biological treatment failures", "presence of comorbidity" and "presence of functional impairment" due to the low frequency with which these were mentioned. In selecting the specific cut-off values for included criteria, we opted to use the most commonly mentioned cut-off values (the mode). Based on the most recent treatment guidelines, for the purpose of this definition we considered SSRIs and SNRIs as current first-line pharmacotherapy options, and CBT current first-line psychotherapy option.[17,19,21,49] Furthermore, the consensus definition for TR-AD requires both a failed pharmacotherapeutic and psychotherapeutic trial, as these were both regularly used as criterion for TR-AD.

This is the first study to systematically assess different criteria for TR-AD. A systematic approach was complicated by the absence of a risk of bias assessment tool for the purpose of the current study. Tools such as the Cochrane risk of bias tool for randomized studies [91] or the RoBANS for non-randomized studies [92] determine the level of confidence with which the *results* of a certain study can be interpreted. However, the data we extracted from studies referred to the definition for treatment resistance they used, not the outcome of the study. Therefore, we chose to assess definition quality by determining the total number of criteria provided and the degree of precision with which the definition was provided. After analyzing these data, it seemed that the quality of included definitions did not impact operationalization of TR-AD.

A limitation in this study was that although integration of the definitions was done systematically the final consensus definition could still reflect some subjective choices by the authors of the current study. Furthermore, it was apparent that some criteria might have been underreported in the included studies, for instance on minimal treatment duration in CBT. Another limitation in our methodology was the lack of studies that incorporated evaluation of pseudo-resistance into their TR-AD definitions. Pseudo-resistance refers to any non-response in treatments that are not employed to their full potential. Before treatment resistance can be deemed present, pseudo-resistance should always be ruled out. In pharmacotherapy trials this could be due to a wrong indication, an inadequate dosage or inadequate duration.[3,35] In psychotherapy trials this could be due to clinicians not following the treatment protocol or patients not being compliant with homework assignments.[3,12] In addition to this, in clinical care it should always be assessed whether the anxiety disorder diagnosis is incorrect, whether another comorbid disorder is the primary problem or whether there are exogenous factors like caffeine overuse, alcohol or substance use or medical diagnoses that contribute to treatment resistance.[3,35] Also, in some studies it was not possible to assess whether previous treatment failures that were counted towards presence of TR-AD consisted of evidence-based anti-anxiety treatments. Finally, although psychological treatments like CBT were repeatedly proven effective in Anxiety Disorders,[14,93] in many parts of the world they are not readily available.[94] Therefore, generalizability of our findings may be limited in these regions.

Furthermore, for the purpose of this study we regarded TR-AD, "refractory anxiety" and other related terms as synonyms. Even though this approach is in line with the majority of the studies, a minority consider TR-AD and "refractory anxiety" to be different entities. For instance, in a Cochrane review, Ipser et al. (2006) propose the term TR for Anxiety Disorder patients who failed one pharmacologic treatment, whereas "refractory anxiety" refers to Anxiety Disorder patients with more than one failed treatment.[95] Their approach can be viewed as a staging approach, distinguishing patients with end-stage TR-AD disorders from those with early stage TR-AD. This approach is also advocated by Cosci & Fava (2013), who propose a staging model for TR Panic Disorders.[44] In their model, the level of TR increases when more treatment regimens within pharmacologic, psychological and combination treatment have failed. In a number of treatment algorithms, a stepped care approach hints to the author's underlying assumption of a staging model for levels of TR.[49] In staging models, treatment decision

making is based on the stage of disease progression in which the patient currently is classified. This could lead to evidence-based stepped-care treatment algorithms. We did not incorporate this staging paradigm for TR-AD into the current paper, as no consensus exists for definitions of TR-AD, nor for staging approaches in TR-AD.

Future studies could empirically investigate the consensus definition for TR-AD. A first step could be to apply the proposed TR-AD definition to an Anxiety Disorder cohort and evaluate the longitudinal course of patients with TR-AD compared to patients without TR-AD. Possibly, this could also yield risk factors for development of TR-AD. Further research could also focus on the validity of a staging approach in TR-AD, as suggested by Cosci & Fava (2013) and Ipser et al. (2006).[44,95]

In depression, a staging paradigm for TR is in use with the Maudsley Staging Method.[23,96,97] A similar approach could be beneficial for Anxiety Disorders. The criteria comprising TR-AD that were described in the current paper could be studied on their merits as individual components in a staging method for TR-AD, to reflect the various degrees of TR-AD.

Conclusions

The majority of studies on treatment resistant Anxiety Disorders (TR-AD) do not demarcate this phenomenon. Across studies that do provide a definition for TR-AD there are many inconsistencies, which are likely to halt progress in Anxiety Disorder research. The current systematic review integrated the current literature into a consensus definition for TR-AD (see Panel 2). This consensus definition should be regarded as a first step to advance the field further. The definition provided in this paper could contribute in harmonization of the process of evaluating presence of TR-AD, which is a necessary first step towards improvement of the prognosis for TR-AD patients.

References

1. Barton S, Karner C, Salih F, Baldwin DS, Edwards SJ. Clinical effectiveness of interventions for treatment-resistant anxiety in older people: A systematic review. *Health Technol Assess (Rockv)*. 2014;18(50):1-62. doi:10.3310/hta18500
2. Chen M-H, Tsai S-J. Treatment-resistant panic disorder: clinical significance, concept and management. *Prog Neuro-Psychopharmacology Biol Psychiatry*. 2016;70:219-226. doi:10.1016/j.pnpbp.2016.02.001
3. Roy-Byrne P. Treatment-refractory anxiety; definition, risk factors, and treatment challenges. *Dialogues Clin Neurosci*. 2015;17(2):191-206.
4. Starcevic V. Treatment of panic disorder: Recent developments and current status. *Expert Rev Neurother*. 2008;8(8):1219-1232. doi:10.1586/14737175.8.8.1219
5. Stein DJ, Seedat S. Unresolved questions about treatment-resistant anxiety disorders. *CNS Spectr*. 2004;9(10):715.
6. Huh J, Goebert D, Takeshita J, Lu BY, Kang M. Treatment of Generalized Anxiety Disorder. *Prim Care Companion CNS Disord*. 2011;13(2). doi:10.4088/PCC.08r00709
7. Bruce SE, Yonkers KA, Otto MW, et al. Influence of Psychiatric Comorbidity on Recovery and Recurrence in Generalized Anxiety Disorder, Social Phobia, and Panic Disorder: A 12-Year Prospective Study. *Am J Psychiatry*. 2005;162(6):1179-1187. doi:10.1176/appi.ajp.162.6.1179
8. Ramsawh HJ, Raffa SD, Edelen MO, Rende R, Keller MB. Anxiety in middle adulthood: effects of age and time on the 14-year course of panic disorder, social phobia and generalized anxiety disorder. *Psychol Med*. 2009;39(4):615-624. doi:10.1017/S0033291708003954
9. Bandelow B, Zohar J, Hollander E, Kasper S, Möller H-J, WFSBP Task force on treatment guidelines for anxiety obsessive-compulsive post-traumatic stress disorders. World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for the pharmacological treatment of anxiety, obsessive-compulsive and post-traumatic stress disorders - First revision. *World J Biol Psychiatry*. 2008;9(4):248-312. doi:10.1080/15622970802465807
10. Bandelow B, Sher L, Bunevicius R, et al. Guidelines for the pharmacological treatment of anxiety disorders, obsessive-compulsive disorder and posttraumatic stress disorder in primary care. *Int J Psychiatry Clin Pract*. 2012;16(2):77-84. doi:10.3109/13651501.2012.667114
11. Farach FJ, Pruitt LD, Jun JJ, Jerud AB, Zoellner LA, Roy-Byrne PP. Pharmacological treatment of anxiety disorders: Current treatments and future directions. *J Anxiety Disord*. 2012;26(8):833-843. doi:10.1016/j.janxdis.2012.07.009

12. Taylor S, Abramowitz JS, McKay D. Non-adherence and non-response in the treatment of anxiety disorders. *J Anxiety Disord*. 2012;26(5):583-589. doi:10.1016/j.janxdis.2012.02.010
13. Jakubovski E, Johnson JA, Nasir M, Müller-Vahl K, Bloch MH. Systematic review and meta-analysis: Dose-response curve of SSRIs and SNRIs in anxiety disorders. *Depress Anxiety*. 2018;(September):1-15. doi:10.1002/da.22854
14. Carpenter JK, Andrews LA, Witcraft SM, Powers MB, Smits JAJ, Hofmann SG. Cognitive behavioral therapy for anxiety and related disorders: A meta-analysis of randomized placebo-controlled trials. *Depress Anxiety*. 2018;35(6):502-514. doi:10.1002/da.22728
15. Tyrer P, Seivewright H, Johnson T. The Nottingham Study of Neurotic Disorder: Predictors of 12-year outcome of dysthymic, panic and generalized anxiety disorder. *Psychol Med*. 2004;34(8):1385-1394. doi:10.1017/S0033291704002569
16. Ipser JC, Carey P, Dhansay Y, Fakier N, Seedat S, Stein DJ. Pharmacotherapy augmentation strategies in treatment-resistant anxiety disorders (Review). *Cochrane Database Syst Rev*. 2006;(4). doi:10.1002/14651858.CD005473.pub2
17. Craske M, Stein MB, Hermann R. Psychotherapy for generalized anxiety disorder in adults. UpToDate. https://www-uptodate-com.vu-nl.idm.oclc.org/contents/psychotherapy-for-generalized-anxiety-disorder-in-adults. Published 2016. Accessed November 22, 2016.
18. Craske M, Bystritsky A, Stein MB, Hermann R. Approach to treating generalized anxiety disorder in adults. UpToDate. https://www.uptodate.com/contents/approach-to-treating-generalized-anxiety-disorder-in-adults. Published 2016. Accessed November 22, 2016.
19. Bystritsky A, Stein MB, Hermann R. Pharmacotherapy for generalized anxiety disorder in adults. UpToDate. https://www-uptodate-com.vu-nl.idm.oclc.org/contents/psychotherapy-for-generalized-anxiety-disorder-in-adults. Published 2016. Accessed November 22, 2016.
20. Baldwin D, Stein MB, Hermann R. Generalized anxiety disorder in adults : Epidemiology , pathogenesis , clinical manifestations, course, assessment, and diagnosis. UpToDate. https://www.uptodate.com/contents/generalized-anxiety-disorder-in-adults-epidemiology-pathogenesis-clinical-manifestations-course-assessment-and-diagnosis. Published 2016. Accessed November 22, 2016.
21. Roy-Byrne PP, Stein MB, Hermann R. Pharmacotherapy for panic disorder in adults. UpToDate. https://www-uptodate-com.vu-nl.idm.oclc.org/contents/pharmacotherapy-for-panic-disorder-in-adults. Published 2016. Accessed November 22, 2016.

22. Howes OD, McCutcheon R, Agid O, et al. Treatment-Resistant Schizophrenia: Treatment Response and Resistance in Psychosis (TRRIP) Working Group Consensus Guidelines on Diagnosis and Terminology. *Am J Psychiatry*. 2016;1(9):appi.ajp.2016.1. doi:10.1176/appi.ajp.2016.16050503
23. Fekadu A, Wooderson S, Donaldson C, et al. A multidimensional tool to quantify treatment resistance in depression: The Maudsley staging method. *J Clin Psychiatry*. 2009;70(2):177-184. doi:10.4088/JCP.08m04309
24. Conway CR, George MS, Sackeim HA. Toward an Evidence-Based, Operational Definition of Treatment-Resistant Depression. *JAMA Psychiatry*. 2017;74(1):9. doi:10.1001/jamapsychiatry.2016.2586
25. Fogelson DL, Leuchter A. Defining Treatment-Resistant Depression. *JAMA Psychiatry*. 2017;74(7):758. doi:10.1001/jamapsychiatry.2017.0837
26. Yoshinaga N, Matsuki S, Niitsu T, et al. Cognitive behavioral therapy for patients with social anxiety disorder who remain symptomatic following antidepressant treatment: A randomized, assessor-blinded, controlled trial. *Psychother Psychosom*. 2016;85(4):208-217. doi:10.1159/000444221
27. Cirillo P, Freire RCR. Pharmacological treatment of panic disorder with non-selective drugs. In: Nardi AE, Freire RCR, Nardi (Ed) AE, Freire (Ed) RCR, eds. *Panic Disorder: Neurobiological and Treatment Aspects.* Freire, Rafael Christophe R.: Springer International Publishing; 2016:289-301. doi:10.1007/978-3-319-12538-1_19
28. Perna G, Caldirola D. Management of Treatment-Resistant Panic Disorder. *Curr Treat Options Psychiatry*. 2017;4(4):371-386. doi:10.1007/s40501-017-0128-7
29. Van Ameringen M, Mancini C. Pharmacotherapy of social anxiety disorder at the turn of the millennium. *Psychiatr Clin North Am*. 2001;24(4):783-803. doi:10.1016/S0193-953X(05)70263-X
30. Stein DJ. Algorithms for primary care: An evidence-based approach to the pharmacotherapy of depression and anxiety disorders. *Prim Psychiatry*. 2004;11(6):55-78.
31. Pollack MH, Smoller JW. Pharmacologic approaches to treatment-resistant panic disorder. In: Pollack MH, Otto MW, Rosenbaum JF, Pollack (Ed) MH, Otto (Ed) MW, Rosenbaum (Ed) JF, eds. *Challenges in Clinical Practice: Pharmacologic and Psychosocial Strategies.* New York, NY, US: Guilford Press; 1996:89-112.
32. Zoun MHH, Koekkoek B, Sinnema H, et al. Effectiveness and cost-effectiveness of a self-management training for patients with chronic and treatment resistant anxiety or depressive disorders: design of a multicenter randomized controlled trial. *BMC Psychiatry*. 2016;16(1). doi:10.1186/s12888-016-0927-1
33. Davies R, Dubovsky S, Gabbert S, Champman M. Treatment resistence in anxiety disorders. *Bull Menninger Clin*. 2000;64(3):A22-A36.

34. Holt RL, Lydiard RB. Management of treatment-resistant panic disorder. *Psychiatry*. 2007;4(10):48-59.
35. Fava GA, Rafanelli C, Tomba E. The Clinical Process in Psychiatry: A Clinimetric Approach. *J Clin Psychiatry*. 2012;73(2):177-184. doi:10.4088/JCP.10r06444
36. Bystritsky A. Treatment-resistant anxiety disorders. *Mol Psychiatry*. 2006;11(9):805-814. doi:10.1038/sj.mp.4001852
37. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med*. 2009;6(7):e1000100. doi:10.1016/j.jclinepi.2009.06.006
38. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*.; 2013.
39. De Salas-Cansado M, Álvarez E, Olivares JM, Carrasco JL, Ferro MB, Rejas J. Modelling the cost-effectiveness of pregabalin versus usual care in daily practice in the treatment of refractory generalised anxiety disorder in Spain. *Soc Psychiatry Psychiatr Epidemiol*. 2013;48(6):985-996. doi:10.1007/s00127-012-0606-6
40. Lohoff FW, Etemad B, Mandos LA, Gallop R, Rickels K. Ziprasidone Treatment of Refractory Generalized Anxiety Disorder. *J Clin Psychopharmacol*. 2010;30(2):185-189. doi:10.1097/JCP.0b013e3181d21951
41. Bakker A, Van Balkom AJLM, Stein DJ. Evidence-based pharmacotherapy of panic disorder. *Int J Neuropsychopharmacol*. 2005;8(3):473-482. doi:10.1017/S1461145705005201
42. Lorenz RA, Jackson CW, Saitz M. Adjunctive use of atypical antipsychotics for treatment-resistant generalized anxiety disorder. *Pharmacotherapy*. 2010;30(9):942-951. doi:10.1592/phco.30.9.942
43. Pollack MH. Refractory Generalized Anxiety Disorder. *J Clin Psychiatry*. 2009;70(suppl 2):32-38. doi:10.4088/JCP.s.7002.06
44. Cosci F, Fava GA. Staging of mental disorders: systematic review. *Psychother Psychosom*. 2013;82(1):20-34. doi:10.1159/000342243
45. Samuel M, Zimovetz EA, Gabriel Z, Beard SM. Efficacy and safety of treatments for refractory generalized anxiety disorder: A systematic review. *Int Clin Psychopharmacol*. 2011;26(2):63-68. doi:10.1097/YIC.0b013e328341bb4a
46. Patterson B, Van Ameringen M. Augmentation strategies for treatment-resistant anxiety disorders: a systematic review and meta-analysis. *Depress Anxiety*. 2016;33(8):728-736. doi:10.1002/da.22525
47. Bandelow B. The medical treatment of obsessive-compulsive disorder and anxiety. *CNS Spectr*. 2008;13(9, 14):37-47.

48. Bandelow B, Zohar J, Hollander E, Kasper S, Moller H-J. World Federation of Societies of Biological Psychiatry (WfSBP) Guidelines for the Pharmacological Treatment of Anxiety, Obsessive-Compulsive and Psttraumatic Stress Disorders. *World J Biol Psychiatry*. 2002;3(904758812):171-199. doi:10.1080/15622970802465807
49. National Institute for Health and Clinical Excellence. *Generalised Anxiety Disorder and Panic Disorder in Adults: Management*. Manchester, United Kingdom; 2011.
50. Stein DJ, Kasper S, Matsunaga H, et al. Pharmacotherapy of social anxiety disorder: an algorithm for primary care - 2001. *Prim Care Psychiatry*. 2001;7(3):107-110. doi:10.1185/135525701317094368
51. Stein DJ, Baldwin DS, Bandelow B, et al. A 2010 evidence-based algorithm for the pharmacotherapy of social anxiety disorder. *Curr Psychiatry Rep*. 2010;12(5):471-477. doi:10.1007/s11920-010-0140-8
52. Stein DJ. Algorithm for the pharmacotherapy of anxiety disorders. *Curr Psychiatry Rep*. 2003;5(4):282-288.
53. Van Ameringen M, Mancini C, Patterson B. Pharmacotherapy for social anxiety disorder and specific phobia. In: Antony MM, Stein MB, Antony (Ed) MM, Stein (Ed) MB, eds. *Oxford Handbook of Anxiety and Related Disorders.* Oxford library of psychology. New York, NY, US: Oxford University Press; 2009:321-333.
54. Deligiannidis KM, Rothschild AJ. Mood and anxiety disorders. In: Rothschild AJ, ed. *The Evidence-Based Guide to Antipsychotic Medications.* 1st ed. Arlington, VA, US: American Psychiatric Publishing, Inc.; 2010:45-100.
55. Baldwin DS, Polkinghorn C. Evidence-based pharmacotherapy of generalized anxiety disorder. In: Stein DJ, Lerer B, Stahl S, Stein (Ed) DJ, Lerer (Ed) B, Stahl (Ed) S, eds. *Evidence-Based Psychopharmacology.* New York, NY, US: Cambridge University Press; 2005:88-104.
56. Gabriel A, Violato C. Adjunctive atomoxetine to SSRIs or SNRIs in the treatment of adult ADHD patients with comorbid partially responsive generalized anxiety (GA): An open-label study. *ADHD Atten Deficit Hyperact Disord*. 2011;3(4):319-326. doi:10.1007/s12402-011-0063-1
57. Kinrys G, Vasconcelos E Sa D, Nery F. Adjunctive zonisamide for treatment refractory anxiety. *Int J Clin Pract*. 2007;61(6):1050-1053. doi:10.1111/j.1742-1241.2007.01365.x
58. Menza MA, Dobkin RD, Marin H. An open-label trial of aripiprazole augmentation for treatment-resistant generalized anxiety disorder [3]. *J Clin Psychopharmacol*. 2007;22(2):207-210. doi:10.1097/01.jcp.0000248620.34541.bc
59. Ociskova M, Prasko J, Latalova K, Kamaradova D, Grambal A. Psychological factors and treatment effectiveness in resistant anxiety disorders in highly comorbid inpatients. *Neuropsychiatr Dis Treat*. 2016;12:1539-1551. doi:10.2147/NDT.S104301

60. Pallanti S, Quercioli L. Resistant social anxiety disorder response to Escitalopram. *Clin Pract Epidemiol Ment Heal*. 2006;2(35). doi:10.1186/1745-0179-2-35
61. Pollack MH, Otto MW, Kaspi SP, Hammerness PG, Rosenbaum JF. Cognitive behavior therapy for treatment-refractory panic disorder. *J Clin Psychiatry*. 1994;55(5):200-205.
62. Simon NM, Hoge EA, Fischmann D, et al. An open-label trial of risperidone augmentation for refractory anxiety disorders. *J Clin Psychiatry*. 2006;67(3):381-385.
63. Snyderman SH, Rynn MA, Rickels K, et al. Open-label Pilot Study of Ziprasidone for Refractory Generalized Anxiety Disorder. *J Clin Psychopharmacol*. 2005;25(5):497-499. doi:10.1097/01.jcp.0000177853.15910.de
64. Solbakken OA, Abbass A. Intensive short-term dynamic residential treatment program for patients with treatment-resistant disorders. *J Affect Disord*. 2015;181:67-77. doi:10.1016/j.jad.2015.04.003
65. Solbakken OA, Abbass A. Symptom- and personality disorder changes in intensive short-term dynamic residential treatment for treatment-resistant anxiety and depressive disorders. *Acta Neuropsychiatr*. 2016;28(05):257-271. doi:10.1017/neu.2016.5
66. Worthington III JJ, Kinrys G, Wygant LE, Pollack MH. Aripiprazole as an augmentor of selective serotonin reuptake inhibitors in depression and anxiety disorder patients. *Int Clin Psychopharmacol*. 2005;20(1):9-11. doi:10.1097/00004850-200501000-00002
67. George MS, Ward HE, Ninan PT, et al. A pilot study of vagus nerve stimulation (VNS) for treatment-resistant anxiety disorders. *Brain Stimul*. 2008;1(2):112-121. doi:10.1016/j.brs.2008.02.001
68. Gabriel A. The mixed amphetamine salt extended release (Adderall XR, Max-XR) as an adjunctive to SSRIs or SNRIs in the treatment of adult ADHD patients with comorbid partially responsive generalized anxiety: an open-label study. *ADHD Atten Deficit Hyperact Disord*. 2010;2(2):87-92. doi:10.1007/s12402-010-0025-z
69. Glue P, Medlicott NJ, Harland S, et al. Ketamine's dose-related effects on anxiety symptoms in patients with treatment refractory anxiety disorders. *J Psychopharmacol*. 2017;31(10):1302-1305. doi:10.1177/0269881117705089
70. Glue P, Neehoff SM, Medlicott NJ, Gray A, Kibby G, McNaughton N. Safety and efficacy of maintenance ketamine treatment in patients with treatment-refractory generalised anxiety and social anxiety disorders. *J Psychopharmacol*. 2018:269881118762073. doi:10.1177/0269881118762073
71. Heldt E, Manfro GG, Kipper L, et al. Treating Medication-Resistant Panic Disorder: Predictors and Outcome of Cognitive-Behavior Therapy in a Brazilian Public Hospital. *Psychother Psychosom*. 2003;72(1):43-48. doi:10.1159/000067188

72. Heldt E, Manfro GG, Kipper L, Blaya C, Isolan L, Otto MW. One-year follow-up of pharmacotherapy-resistant patients with panic disorder treated with cognitive-behavior therapy: Outcome and predictors of remission. *Behav Res Ther*. 2006;44(5):657-665. doi:10.1016/j.brat.2005.05.003
73. Hoge EA, Worthington III JJ, Kaufman RE, et al. Aripiprazole as augmentation treatment for refractory generalized anxiety disorder and panic disorder. *CNS Spectr*. 2008;13(6):522-527.
74. Hollifield M, Thompson PM, Ruiz JE, Uhlenhuth EH. Potential effectiveness and safety of olanzapine in refractory panic disorder. *Depress Anxiety*. 2005;21(1):33-40. doi:10.1002/da.20050
75. Katzman MA, Vermani M, Jacobs L, et al. Quetiapine as an adjunctive pharmacotherapy for the treatment of non-remitting generalized anxiety disorder: A flexible-dose, open-label pilot trial. *J Anxiety Disord*. 2008;22(8):1480-1486. doi:10.1016/j.janxdis.2008.03.002
76. Brawman-Mintzer O, Knapp RG, Nietert PJ. Adjunctive risperidone in generalized anxiety disorder: A double-blind, placebo-controlled study. *J Clin Psychiatry*. 2005;66(10):1321-1325.
77. Castle C, Gray A, Neehoff S, Glue P. Effect of ketamine dose on self-rated dissociation in patients with treatment refractory anxiety disorders. *J Psychopharmacol*. 2017;31(10):1306-1311. doi:10.1177/0269881117725685
78. Gloster AT, Sonntag R, Hoyer J, et al. Treating treatment-resistant patients with panic disorder and agoraphobia using psychotherapy: A randomized controlled switching trial. *Psychother Psychosom*. 2015;84(2):100-109. doi:10.1159/000370162
79. Hirschmann S, Dannon PN, Iancu I, Dolberg OT, Zohar J, Grunhaus L. Pindolol Augmentation in Patients With Treatment-Resistant Panic Disorder: A Double-Blind, Placebo-Controlled Trial. *J Clin Psychopharmacol*. 2000;20(5):556-559. doi:10.1097/00004714-200010000-00011
80. Pollack MH, Simon NM, Zalta AK, et al. Olanzapine augmentation of fluoxetine for refractory generalized anxiety disorder: A placebo controlled study. *Biol Psychiatry*. 2006;59(3):211-215. doi:10.1016/j.biopsych.2005.07.005
81. Rickels K, Shiovitz TM, Ramey TS, Weaver JJ, Knapp LE, Miceli JJ. Adjunctive therapy with pregabalin in generalized anxiety disorder patients with partial response to SSRI or SNRI treatment. *Int Clin Psychopharmacol*. 2012;27(3):142-150. doi:10.1097/YIC.0b013e328350b133
82. Simon NM, Otto MW, Worthington JJ, et al. Next-Step Strategies for Panic Disorder Refractory to Initial Pharmacotherapy. *J Clin Psychiatry*. 2009;70(11):1563-1570. doi:10.4088/JCP.08m04485blu

83. Bakish D, Hooper CL, West DL, Miller C, Blanchard A, Bashir F. Moclobemide and specific serotonin re-uptake inhibitor combination treatment of resistant anxiety and depressive disorders. *Hum Psychopharmacol Clin Exp*. 1995;10(2):105-109. doi:10.1002/hup.470100205
84. Cowley DS, Ha EH, Roy-Byrne PP. Determinants of pharmacologic treatment failure in panic disorder. *J Clin Psychiatry*. 1997;58(12):555-561.
85. Durham RC, Higgins C, Chambers JA, Swan JS, Dow MGT. Long-term outcome of eight clinical trials of CBT for anxiety disorders: Symptom profile of sustained recovery and treatment-resistant groups. *J Affect Disord*. 2012;136(3):875-881. doi:10.1016/j.jad.2011.09.017
86. Kinrys G, Worthington JJ, Wygant L, Nery F, Reese H, Pollack MH. Levetiracetam as adjunctive therapy for refractory anxiety disorders. *J Clin Psychiatry*. 2007;68(7):1010-1013.
87. Milrod B, Altemus M, Gross C, et al. Adult separation anxiety in treatment nonresponders with anxiety disorders: Delineation of the syndrome and exploration of attachment-based psychotherapy and biomarkers. *Compr Psychiatry*. 2016;66:139-145. doi:10.1016/j.comppsych.2016.01.004
88. Aarre TF. Phenelzine efficacy in refractory social anxiety disorder: A case series. *Nord J Psychiatry*. 2003;57(4):313-315. doi:10.1080/08039480310002110
89. Otto MW, Pollack MH, Penava SJ, Zucker BG. Group cognitive-behavior therapy for patients failing to respond to pharmacology for panic disorder: A clinical case series. *Behav Res Ther*. 1999;37(8):763-770. doi:10.1016/S0005-7967(98)00176-4
90. Tesar GE, Rosenbaum JF. Successful use of clonazepam in patients with treatment-resistant panic disorder. *J Nerv Ment Dis*. 1986;174(8):477-482.
91. Higgins J, Green S. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [Updated March 2011]*. The Cochrane Collaboration.; 2011.
92. Kim SY, Park JE, Lee YJ, et al. Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. *J Clin Epidemiol*. 2013;66(4):408-414. doi:10.1016/j.jclinepi.2012.09.016
93. Bandelow B, Reitt M, Röver C, Michaelis S, Görlich Y, Wedekind D. Efficacy of treatments for anxiety disorders: A meta-analysis. *Int Clin Psychopharmacol*. 2015;30(4):183-192. doi:10.1097/YIC.0000000000000078
94. Saxena S, Thornicroft G, Knapp M, Whiteford H. Resources for mental health: scarcity, inequity, and inefficiency. *Lancet*. 2007;370(9590):878-889. doi:10.1016/S0140-6736(07)61239-2
95. Ipser JC, Carey P, Dhansay Y, Fakier N, Seedat S, Stein DJ. Pharmacotherapy augmentation strategies in treatment- resistant anxiety disorders. *Cochrane Database Syst Rev*. 2006;(4). doi:10.1002/14651858.

96. Peeters FPML, Ruhe HG, Wichers M, et al. The Dutch Measure for quantification of Treatment Resistance in Depression (DM-TRD): an extension of the Maudsley Staging Method. *J Affect Disord*. 2016;205:365-371. doi:10.1016/j.jad.2016.08.019

97. van Belkum SM, Geugies H, Lysen TS, et al. Validity of the Maudsley Staging Method in Predicting Treatment-Resistant Depression Outcome Using the Netherlands Study of Depression and Anxiety. *J Clin Psychiatry*. 2018;79(1). doi:10.1103/PhysRevLett.113.113903

Supplementary materials.

Abbreviations used in eTables

ACT	Acceptance and Commitment Therapy
ADs	antidepressants
BDZ	Benzodiazepine
CAS+PA	Clinical Anxiety Scale with panic attacks
CAU	care as usual
CBT	Cognitive Behavioral Therapy
CGI-I	Clinical Global Impression Improvement Scale
CGI-S	Clinical Global Impression Severity Scale
DSM-IV	Diagnostic and Statistical Manual, fourth edition
ECT	Electroconvulsive therapy
ER	extended release
FU	follow-up
GAD	Generalized Anxiety Disorder
GAF	Global Assessment of Functioning
HAM-A	Hamilton Anxiety Rating Scale
HAM-D	Hamilton Depression Rating Scale
H	high
IIP-64	Inventory of Interpersonal Problems
L	low
LSAS	Leibowitz Social Anxiety Scale
M	medium
MAO-I	Monoamine Oxidase inhibitor
MDD	Major Depressive Disorder
MGHAP CGI-S	Massachusetts General Hospital Anchored Panic CGI-S
MI	Mobility Inventory
na	not available
OCD	Obsessive Compulsive Disorder
OQ-45.2	outcome measure 45.2
PAS	Panic and Agoraphobia Scale
PD	Panic Disorder (with or without agoraphobia)
PDSS	Panic Disorder Severity Scale
PSQ	Panic Self Questionnaire
PTSD	Post Traumatic Stress Disorder
RCT	randomized controlled trial
refr	refractory
SAD	Social Anxiety Disorder
SCL-90-R	Symptom Checklist-90, Revised
SDS	Sheehan Disability Scale
SNRI	selective Serotonin and Norepinephrine Reuptake Inhibitor
SP	Specific Phobia
SRI	Serotonin Reuptake Inhibitor
SSRI	Selective Serotonin Reuptake Inhibitor
TCA	Tricyclic Antidepressant
TR	treatment resistant
XR	extended release

eTable 1. Study characteristics of included trials, cohort studies and meta-analyses.

Authors, year of publication	Study design	No of studies/ subjects	Population	Intervention	Comparator	Follow-up (FU) duration	Primary outcome	Conclusions
Aarre, 2003	Case series	n=7	Refractory SAD	Phenelzine (15-90 mg twice daily)	None	unspecified	Anxiety severity	FU > baseline
Bakish et al., 1995	Retrospective cohort	n=31	Resistant PD, MDD, or Dysthymia	Moclobemide (35-800 mg/day) and SSRI (1-300 mg/day) combination treatment	None	unspecified	Symptom severity	FU > baseline
Barton, Karner, Salih, Baldwin, & Edwards, 2014	Systematic review, meta-analysis	0 studies included	TR GAD, PD, SAD, SP, OCD and PTSD in older adults	Any interventions for TR	Placebo, no intervention, or another active intervention	unspecified	Anxiety severity	no evidence base for treatments of TR anxiety in older adults
Brawman-Mintzer, Knapp, & Nietert, 2005	RCT	n=40	Refractory GAD	Adjunctive risperidone (0.5-1.5 mg/ day)	Placebo	5 weeks	Anxiety severity	Risperidone > placebo
Castle, Gray, Neehoff, & Glue, 2017	Combined open-label-trial/ RCT	*N=18*	TR GAD and/or SAD	Ketamine in 0.25 mg/kg, 0.5 mg/kg, and 1 mg/kg	midazolam 0.01 mg/kg	7 days	self-reported levels of dissociation	Intervention induces more dissociative symptoms
Cowley, Ha, & Roy-Byrne, 1997	Retrospective cohort	n=106	TR PD	No intervention	None	Not specified	Predictors of treatment failure	Predictors: Medication intolerance, inadequate dose or duration
De Salas-Cansado et al., 2013	Cost-effectiveness study	n=902	Refractory GAD	Pregabalin mono- or adjunctive therapy (average dose 218 mg/day)	CAU	6 months	Cost-effectiveness	Intervention > CAU
Durham, Higgins, Chambers, Swan, & Dow, 2012	Follow-up after retrospective cohort	n=336	CBT study participants with GAD, PD, PTSD or MDD	Adjunctive CBT	None	2-14 years post-treatment	Severity of anxiety, functional status, healthcare usage	38% recovered with minimal treatment, 30% poor outcome despite treatment
Gabriel, 2010	Open-label trial	n=32	Refractory GAD	Adjunctive Adderall XR (5-50 mg/day)	None	12 weeks	Anxiety severity	FU > baseline
Gabriel & Violato, 2011	Open-label trial	n=29	Refractory GAD	Adjunctive atomoxetine (10-40 mg twice daily)	None	12 weeks	Anxiety severity	FU > baseline
George et al., 2008	Open-label trial	n=11	TR PD, OCD and PTSD	Vagus nerve stimulation	None	4.5 years	Symptom severity	FU > baseline
Glue et al., 2017	Open-label trial	*N=12*	TR GAD and/or SAD	Ketamine in 0.25 mg/kg, 0.5 mg/kg, and 1 mg/kg	none	7 days	anxiety severity	Dose response relation present

Authors, year of publication	Study design	No of studies/ subjects	Population	Intervention	Comparator	Follow-up (FU) duration	Primary outcome	Conclusions
Glue et al., 2018	Open-label trial	*N=20*	TR GAD and/or SAD who responded in ketamine trial	Ketamine maintenance (3 months)	none	3 months post intervention	- Anxiety severity - tolerability - social and work functioning	Maintenance treatment sustains remission
Gloster et al., 2015	RCT	n=43	TR PD	Adjunctive acceptance and commitment therapy (ACT)	Waiting-list	6 months	Symptom severity	ACT > waiting-list
Heldt et al., 2003	Open-label trial	n=32	TR PD	Adjunctive CBT	None	12 weeks	1: Symptomatic change 2: negative outcome predictors	1: FU > baseline 2: Predictors: depression, neurotic defense style
Heldt et al., 2006	Open-label trial	n=64	TR PD	Adjunctive CBT	None	12 weeks + 1-year post-treatment	1: Anxiety severity 2: Negative outcome predictors	1: FU > baseline 2: Predictors: comorbid dysthymia, SP, GAD
Hirschmann et al., 2000	RCT	n=26	TR PD	Adjunctive pindolol (2.5 mg three times daily)	Placebo	4 weeks	Anxiety severity	Intervention > placebo
Hoge et al., 2008	Open-label trial	n=23	Refractory GAD and PD	Adjunctive aripiprazole (2.5- 30 mg/ day)	None	8 weeks	Anxiety severity	FU > baseline
Hollifield, Thompson, Ruiz, & Uhlenhuth, 2005	Open-label trial	n=10	Refractory PD	Olanzapine (2.5-20 mg/day) after tapering off current medication	None	8 weeks	Anxiety severity, impairment and safety	FU > baseline
Ipser et al., 2006	Systematic review and meta-analysis	28 RCTs	TR GAD, PD, SAD, OCD and PTSD	Pharmacotherapeutic augmentation	Placebo or other medication (e.g. monotherapy)	4-20 weeks	Response	Intervention > control
Katzman et al., 2008	Open-label trial	n=40	TR GAD	Adjunctive quetiapine (25-800 mg/ day)	None	12 weeks	Anxiety severity	FU > baseline
Kinrys, Vasconcelos E Sa, & Nery, 2007	Open-label trial	n=10	Refractory GAD, PD, SAD and PTSD	Adjunctive zonisamide (100-300 mg/day)	None	4-20 weeks	Anxiety severity	FU > baseline
Kinrys et al., 2007	Retrospective cohort	n=40	Refractory GAD, PD, SAD and PTSD	Adjunctive levetiracetam (250-3000 mg/day)	None	9.3 ± 5.1 weeks	Anxiety severity	FU > baseline
Lohoff, Etemad, Mandos, Gallop, & Rickels, 2010r	RCT	n=62	TR GAD	Adjunctive ziprasidone (20-80 mg/ day)	Placebo	8 weeks	Anxiety severity	Intervention = placebo
Menza, Dobkin, & Marin, 2007	Open-label trial	n=9	TR GAD	Adjunctive aripiprazole	None	6 weeks	Anxiety severity and improvement	FU > baseline

Authors, year of publication	Study design	No of studies/ subjects	Population	Intervention	Comparator	Follow-up (FU) duration	Primary outcome	Conclusions
Milrod et al., 2016	Prospective cohort	n=46	Refractory anxiety disorders	Adjunctive Panic Focused Psychodynamic Psychotherapy-eXtended Range	None	12 weeks	1: Symptomatic change after FU 2: Separation anxiety prevalence 3: Biomarker levels	1: FU > baseline 2: 80% had separation anxiety 3: no significant change
Ociskova, Prasko, Latalova, Kamaradova, & Grambal, 2016	Open-label trial	n=109	TR GAD, PD, SP and mixed anxiety and depressive disorder	Inpatient CBT + pharmacotherapy	Inpatient psychodynamic therapy + pharmacotherapy	6 weeks	Psychological predictors of treatment outcome	Multiple predictors found
Otto, Pollack, Penava, & Zucker, 1999a	Case series	n=24	TR PD	Adjunctive CBT	None	12 weeks	Anxiety severity	FU > baseline
Pallanti & Quercioli, 2006	Open-label trial	n=29	TR SAD	Escitalopram (10-20 mg/day) after wash-out	None	12 weeks	Anxiety severity	FU > baseline
Patterson & Van Ameringen, 2016	Systematic review and meta-analysis	6 RCTs, n=557	TR GAD, PD and SAD	Pharmacotherapeutic or CBT augmentation to SRI treatment	Placebo	6-12 weeks	Clinical improvement	Intervention = placebo
Pollack, Otto, Kaspi, Hammerness, & Rosenbaum, 1994	Open-label trial	n=15	Refractory PD	Adjunctive CBT	None	12 weeks + 1-8 months post-treatment	Anxiety severity	FU > baseline
Pollack et al., 2006	RCT	n=21	TR GAD	Adjunctive olanzapine (2.5-20 mg/day)	Placebo	6 weeks	Anxiety severity	Intervention > placebo
Rickels et al., 2012	RCT	n=356	Refractory GAD	Adjunctive pregabalin (150-600 mg/day)	Adjunctive placebo	8 weeks	Anxiety severity	Intervention > placebo
Simon et al., 2006social anxiety disorder (SAD	Open-label trial	n=30	Refractory GAD, PD and SAD	Adjunctive risperidone (0.25-3 mg/day)	None	8 weeks	Anxiety severity	FU > baseline
Simon et al., 2009	Combined RCT/ Open-label trial	n=46	Refractory PD	Phase 1: Prospective assessment of refractoriness Phase 2: SSRI dose optimization Phase 3: Adjunctive CBT	Phase 1: No control arm Phase 2: placebo Phase 3: medication optimization	24 weeks + 3 months post-treatment	Anxiety severity	Phase 2 and 3: Both groups improved, Intervention = placebo
Snyderman et al., 2005	Open-label trial	n=13	TR GAD	Ziprasidone (20-80 mg/day)	None	7 weeks	Anxiety severity	FU > baseline
Solbakken & Abbass, 2015	Open-label trial	n=60	TR anxiety or depressive disorder	Intensive inpatient treatment program[1]	None	8 weeks + 14 months post-treatment	Symptom severity	FU > baseline

Authors, year of publication	Study design	No of studies/ subjects	Population	Intervention	Comparator	Follow-up (FU) duration	Primary outcome	Conclusions
Solbakken & Abbass, 2016	Open-label trial	n=95	TR anxiety or depressive disorder	Intensive inpatient treatment program[1]	None	8 weeks + 14 months post-treatment	Symptom severity	FU > baseline
Tesar & Rosenbaum, 1986	Case series	n=10	TR PD	Adjunctive clonazepam (1.5-8 mg/day)	None	8-68 weeks	Symptom improvement	FU > baseline
Worthington III, Kinrys, Wygant, & Pollack, 200	Open-label trial	n=17	TR anxiety disorder or MDD	Adjunctive aripiprazole (7.5-30 mg/ day)	None	12 weeks	Symptom severity	FU > baseline
Yoshinaga et al., 2016	Assessor blinded open-label trial	n=42	TR SAD	Adjunctive CBT	CAU	16 weeks	Anxiety severity	Intervention > CAU
Zoun et al., 2016	Trial protocol	*na*	TR anxiety and depressive disorders	Self-management for Chronic Anxiety and Depression	CAU	18 months	1: Quality of life 2: Symptom severity 3: Costs	Not yet available

[1] the program consisted of individual psychotherapy, group psychotherapy, psychopharmacological treatment, and therapeutic group activities.

eTable 2. Study characteristics for included reviews, treatment guidelines and book chapters.

Authors, year of publication	Study design	Population	Study description
Bakker, Van Balkom, & Stein, 2005	Narrative review	Treatment refractory PD	Examines first-line pharmacotherapy, optimal duration of maintenance pharmacotherapy, and optimal approach to treatment refractoriness in PD patients.
Baldwin & Polkinghorn, 2005	Book chapter	Non-responsive GAD	Examines first-line pharmacotherapy, optimal duration of treatment, and best interventions after non-response of first-line and second-line treatments in GAD patients
Bandelow, Zohar, Hollander, Kasper, & Moller, 2002	Guideline	GAD, PD, SAD, SP, OCD and PTSD	Guideline for the pharmacological treatment of GAD, PD, SAD, SP, OCD and PTSD
Bandelow, 2008	Guideline	GAD, PD, SAD, SP, OCD and PTSD	First revision of a guideline for the pharmacological treatment of GAD, PD, SAD, SP, OCD and PTSD
Bandelow et al., 2008	Narrative review	GAD, PD, SAD and OCD	Summary of pharmacological treatment recommendations for GAD, PD, SAD and OCD
Bystritsky, 2006	Narrative review	TR GAD, PD, SAD and OCD	Reviews reasons for TR and strategies for improving outcome in TR patients
Chen & Tsai, 2016	Narrative review	TR PD	Presents definitions, risk factors, pathophysiology hypotheses and therapeutic strategies for TR PD
Cosci & Fava, 2013	Systematic review	PD	Synthesizes 78 studies to describe the different models of staging currently known in clinical psychology and psychiatry (including a staging model for panic disorder)
Deligiannidis & Rothschild, 2010	Book chapter	GAD, OCD, PTSD, bipolar disorder, depressive disorders	Reviews the evidence base for the use of antipsychotic medication in this population
Holt & Lydiard, 2007	Narrative review	PD	Reviews the pathophysiology, existing and emerging treatment options and strategies for optimizing treatment response in PD and presents a diagnostic approach for unresponsive PD
Lorenz, Jackson, & Saitz, 2010	Narrative review	TR GAD	Reviews the safety and efficacy of atypical antipsychotics as augmentation to pharmacotherapy in TR GAD patients
National Institute for Health and Clinical Excellence, 2011	Guideline	GAD and PD	Treatment guideline for PD and GAD patients
Pollack, 2009	Narrative review	GAD	Summarizes clinical and demographic characteristics and pharmacotherapeutic strategies for GAD
Samuel, Zimovetz, Gabriel, & Beard, 2011	Systematic Review	GAD	Reviews eight studies with regard to efficacy and safety of treatments for refractory GAD
Starcevic, 2008	Narrative review	PD	Reviews developments and future challenges in the treatment of PD
Stein et al., 2001	Algorithm	SAD	Primary care pharmacotherapy algorithm for SAD
Stein et al., 2010	Algorithm	SAD	Updated version of a primary care pharmacotherapy algorithm for SAD

Authors, year of publication	Study design	Population	Study description
Stein, 2003	Algorithm	GAD, PD, SAD, OCD and PTSD	Pharmacotherapy algorithm for GAD, PD, SAD, OCD and PTSD
Stein, 2004	Algorithm	GAD, PD, SAD, OCD, PTSD, MDD	Primary care pharmacotherapy algorithm for MDD, GAD, PD, SAD, PTSD and OCD
Van Ameringen, Mancini, & Patterson, 2009	Book chapter	SAD and SP	Overview of efficacy of pharmacological treatments for SAD and SP

eTable 3. Definitions for treatment resistance in anxiety disorders, as used by authors.

Authors, year of publication	Population	Degree of precision	Treatment resistance definition
Aarre, 2003	Refractory SAD	Medium	"failing to respond to other treatments that have been proved effective in this condition"
Bakish et al., 1995	Resistant PD, MDD, or Dysthymia	Medium	"conventional treatment for their psychiatric disorder failed"
Bakker et al., 2005in order to address the questions (1	Treatment refractory PD	Medium	"not all patients respond to the first trial of medication"
Baldwin & Polkinghorn, 2005	Non-responsive GAD	Low	"patients who have not responded to first-line or second-line treatments"
Bandelow et al., 2002	GAD, PD, SAD, SP, OCD and PTSD	Medium	"When initial treatment fails" PD: "patients (...) who were resistant to several antipanic drug treatments" SAD: "partial response to a SSRI", "non-responders to SSRIs"
Bandelow, 2008	GAD, PD, SAD, SP, OCD and PTSD	High	"do not fulfill response criteria after initial standard treatment", "While no universally accepted criteria exist, a commonly used threshold for response is a >50% improvement in the total score of a ...) rating scale"
Bandelow et al., 2008	GAD, PD, SAD and OCD	High	"(...) do not fulfill response criteria after initial standard treatment. While no universally accepted criteria exist, a commonly used threshold for response is a 50% improvement in the total score of a (...) rating scale" PD: "When initial treatments have failed", "patients having residual symptoms despite being on an adequate dose of medication", "patients who responded insufficiently to CBT" SAD: "non-responsive to standard treatment" GAD: "patients remaining symptomatic despite initial anxiolytic treatment"
Barton et al., 2014	TR GAD, PD, SAD, SP, OCD and PTSD in older adults	High	"no evidence of substantial improvement after 4 weeks' treatment with a treatment for which there is evidence of clinical effectiveness in the treatment of anxiety"
Brawman-Mintzer et al., 2005	Refractory GAD	Medium	"patients who remain symptomatic despite ongoing standard anxiolytic treatment"
Bystritsky, 2006	TR GAD, PD, SAD and OCD	High	"no restoration or near restoration of functional status in the presence (absence) of tolerable treatment"
Castle et al., 2017	TR GAD and/or SAD	Medium	"HAM-A[1] score of >20, and/or a (...) LSAS[1] score of >60"
Chen & Tsai, 2016	TR PD	High	"the failure to achieve either of two remission criteria after at least 6 months of optimal treatment" Remission criteria 1: complete resolution in 5 principal domains: panic attacks (...), anticipatory anxiety, panic-related phobias, well-being/severity of illness, and functional and social impairment caused by the panic disorder Remission criteria 2: (...) PDSS ≤ 3, HAM-D ≤ 7

Authors, year of publication	Population	Degree of precision	Treatment resistance definition
Cosci & Fava, 2013	PD	High	Stage 1: a pharmacological/psychological intervention fails to give benefits Stage 2: the failure involves 2 different interventions, including at least 1 involving psychotherapeutic treatment Stage 3: involves the failure of 3 or more adequate therapeutic interventions, including at least 1 involving Psychotherapy Stage 4: there is failure of 3 or more adequate therapeutic trials, including at least 1 combination of psychotherapy and pharmacotherapy."
Cowley et al., 1997	TR PD	Medium	TR: "patients [who] fail initial treatment" True TR: "failure to respond to an effective medication given at a therapeutic dose for an appropriate length of time"
Deligiannidis & Rothschild, 2010	GAD, OCD, PTSD, bipolar disorder, depressive disorders	Medium	"partial and non-responders to the first-line GAD therapies"
De Salas-Cansado et al., 2013	Refractory GAD	High	"persistent symptoms/suboptimal response (...) after a standard dose regimen of any anti-anxiety drug, alone or in combination, for at least 6 months"
Durham et al., 2012	CBT study participants with GAD, PD, PTSD or MDD	Medium	"a clinical diagnosis of at least one anxiety disorder despite receiving either a moderate amount or a lot of interim treatment", "with either medication or psychological therapy"
Gabriel, 2010	Refractory GAD	Medium	"failed to respond to at least one 8-week trial of SSRI, or SNRI (...) with > 50% reduction in anxiety symptoms from baseline"
Gabriel & Violato, 2011	Refractory GAD	Medium	"partial response to at least one 8-week trial of SSRI or SNRI"
George et al., 2008	TR PD, OCD and PTSD	High	"tried and failed pharmacotherapy and psychotherapy"
Glue et al., 2017	TR GAD and/or SAD	Medium	"(...) HAM-A score of >20, and/or a (...) LSAS score of >60"
Glue et al., 2018	TR GAD and/or SAD who responded in ketamine trial	Medium	"Hamilton Anxiety Scale (...) score of >20, and/or a Liebowitz Social Anxiety Scale (...) score of >60"
Gloster et al., 2015	TR PD	Medium	"have had one or more previous courses of psychological and/or pharmacological treatment consistent with state-of-the-art practice"
Heldt et al., 2003	TR PD	Medium	"the patients must have had residual symptoms of panic attacks, anticipatory anxiety and phobic avoidance despite having been on an adequate and stable dose of an SSRI for at least 4 months"
Heldt et al., 2006	TR PD	Medium	"patients had to have residual symptoms of PD such as panic attacks, anticipatory anxiety and phobic avoidance despite being on a stable dose of medications for at least 4 months"
Hirschmann et al., 2000	TR PD	Medium	"less than 20% reduction in score (...) after an 8-week trial of fluoxetine 20mg/day"
Hoge et al., 2008	Refractory GAD and PD	Medium	"DSM-IV criteria for GAD or PD despite initial pharmacotherapy with an adequate (or highest tolerated) dose of an anxiolytic agent, including an SSRI or a (...) SNRI and/or BDZs or trazodone or bupropion, initiated at least 8 weeks prior to study initiation"
Hollifield et al., 2005	Refractory PD	High	"having failed two adequate therapeutic trials of either medication or (...) CBT"

Authors, year of publication	Population	Degree of precision	Treatment resistance definition
Holt & Lydiard, 2007	PD	High	"an inadequate response to what is generally considered adequate treatment"
Ipser et al., 2006	TR GAD, PD, SAD, OCD and PTSD	High	TR: "fail to respond to a first-line intervention" treatment refractory: "no change or whose symptoms worsen following several different interventions will be referred to as treatment-refractory"
Katzman et al., 2008	TR GAD	Medium	"not responding to (...) a traditional therapy (...) a minimum of 8 weeks' treatment with an appropriate dose of a traditional anxiolytic"
Kinrys et al., 2007	Refractory GAD, PD, SAD and PTSD	Medium	"history of persistent anxiety despite adequate anxiolytic treatment (defined as 8 weeks or more of treatment (...)"
Kinrys et al., 2007	Refractory GAD, PD, SAD and PTSD	Medium	"did not have a full response during their current episode despite an adequate dose and duration of their anxiolytic trial, defined as persistence of anxiety symptoms (...)"
Lohoff et al., 2010	TR GAD	Medium	"treatment failure of at least 1 adequate trial of an SSRI, an SNRI, a BDZ, or a combination of these agents"
Lorenz et al., 2010	TR GAD	High	"not respond to at least one antidepressant at an adequate dose for an adequate duration"
Menza et al., 2007	TR GAD	Medium	"a (...) HAM-A score of 14 or greater, and a (...) CGI-S score of at least 4 (moderately ill). (...) patients had been treated with an antidepressant at a therapeutic dose (fluoxetine, 40 mg; paroxetine, 40 mg; escitalopram, 20 mg; bupropion, 250mg; venlafaxine, 150 mg; and mirtazapine, 30 mg) for an adequate duration (at least 6 weeks)."
Milrod et al., 2016	Refractory anxiety disorders	Medium	"prominent anxiety (...) and could provide evidence of nonresponse (clinically persistent anxiety) to at least one evidence-based antianxiety treatment, specifically: (a) at least 2 months of an adequately dosed SSRI or TCA equivalent (...) or (b) ≥11 weeks of CBT"
National Institute for Health and Clinical Excellence, 2011	GAD and PD	Medium	"severe anxiety with marked functional impairment in conjunction with: a risk of self-harm or suicide or significant comorbidity, such as substance misuse, personality disorder or complex physical health problems or self-neglect or an inadequate response to step 3 interventions."
Ociskova et al., 2016	TR GAD, PD, SP and mixed anxiety and depressive disorder	Medium	"patients were resistant to outpatient pharmacological and psychotherapeutic treatment (i.e., they were chronically unresponsive to standard outpatient pharmacological and/or psychological treatment)"
Otto et al., 1999	TR PD	Medium	"Current CGI-S ≥ 4 with a minimum of 2 months of treatment with any of the following (...) (1) Any tricyclic (...) (2) Fluoxetine (...) (3) Sertraline (...) (4) Paroxetine (...) (5) Nefazodone (...) (6) Phenelzine (...) (7) Alprazolam (...) (8) Clonazepam (...)"
Pallanti & Quercioli, 2006	TR SAD	Medium	"failed at least one adequate trial of paroxetine treatment (...) 'Failure' was defined as: 1: experiencing less than a 35% decrease from baseline (...), and a score of 'minimal improvement' or less on the (...) CGI-I after 12 weeks of treatment 2: having stopped the medication in the first 3 weeks of treatment because of intolerable side effects or lack of compliance"
Patterson & Van Ameringen, 2016	TR GAD, PD and SAD	Medium	"either a less than 50% improvement in the total score of a commonly used anxiety rating scale or a nonresponse to an adequate dose of first-line pharmacological treatment of a SRIs for 4–5 weeks."

Authors, year of publication	Population	Degree of precision	Treatment resistance definition
Pollack et al., 1994	Refractory PD	Medium	Treatment refractory: "incomplete response to an adequate trial of medication" Incomplete response to pharmacotherapy: "the continued presence, despite at least 3 months of medication treatment, of panic attacks, phobic fear or avoidance, or anticipatory anxiety severe enough to interfere with function."
Pollack et al., 2006	TR GAD	Medium	"remain symptomatic despite treatment with standard agents such as the SSRIs"
Pollack, 20099	GAD	Medium	"remaining symptomatic or to not respond at all to first-line medication"
Rickels et al., 2012	Refractory GAD	Medium	"partial response to 8 weeks of (...) treatment with venlafaxine-extended release (XR), escitalopram, or paroxetine, and who had a suboptimal response (by history) to at least one previous treatment for GAD."
Samuel et al., 2011	GAD	Medium	"failed to respond adequately to at least one earlier treatment for GAD."
Simon et al., 2006	Refractory GAD, PD and SAD	Medium	"significant persistent symptoms despite initial pharmacotherapy (...) an adequate (or maximally tolerated) dose of an established anxiolytic (SSRI or SNRI and/or (...) BDZs) initiated at least 8 weeks prior"
Simon et al., 2009	Refractory PD	Medium	"lack of remission"
Snyderman et al., 2005	TR GAD	Medium	"a score of 4 or greater on the (...) CGI-S, a (...) HAM-A score 16 or greater after 8 weeks of treatment with at least 1 first-line antianxiety agent"
Solbakken & Abbass, 2015	TR anxiety or depressive disorder	Medium	"failure to respond with symptomatic relief and improved occupational or interpersonal functioning to three or more prior attempts at treatment for the ongoing psychiatric disorder. The previous treatment attempts could be either medication efforts or psychotherapeutic/psychosocial efforts, or most commonly a combination of both"
Solbakken & Abbass, 2016	TR anxiety or depressive disorder	Medium	"failure to respond with symptomatic relief and improved occupational or interpersonal functioning despite three or more prior attempts at treatment for the ongoing disorder. The previous treatment attempts could be either medication efforts or psychotherapeutic/psychosocial efforts, or most commonly a combination of both"
Starcevic, 2008	PD	High	"a patient with PD can be considered resistant to pharmacological treatment if he/she has not fully responded to an adequate treatment with two SSRIs, venlafaxine ER, one of the TCAs (imipramine or clomipramine), a high-potency BDZs, or a combination of one of these antidepressants with a BDZ"
Stein et al., 2001	SAD	Low	"who have failed to respond to a (...) SSRI"
Stein et al., 2010	SAD	Medium	"fail to respond to treatment with the first SSRI/SNRI"
Stein, 2003	GAD, PD, SAD, OCD and PTSD	Medium	"failure of an adequate clinical trial of medication"
Stein, 2004	GAD, PD, SAD, OCD, PTSD, MDD	Low	"failure of an adequate clinical trial of medication"
Tesar & Rosenbaum, 1986	TR PD	Low	"failure to respond to, or refusal to be treated with, each of the major antipanic agents, including a TCA, a MAO-i, and alprazolam. Failure to respond is defined by presence of one or more of the following: a) lack of therapeutic effect, b) tolerance to drug effect, c) suboptimal therapeutic effect, or d) medication intolerance due to side effects"

Authors, year of publication	Population	Degree of precision	Treatment resistance definition
Van Ameringen et al., 2009	SAD and SP	Low	"not respond, or has a partial response to a standard, first line treatment"
Worthington III et al., 2005	TR anxiety disorder or MDD	Medium	"did not have a full response during their current episode despite an adequate dose and duration of their SSRI trial"
Yoshinaga et al., 2016	TR SAD	Medium	" 'patients who remain symptomatic following antidepressant treatment' "
Zoun et al., 2016	TR anxiety and depressive disorders	Medium	"prolonged treatment in a specialized outpatient mental health service according to the professional is unlikely to improve clinical outcomes (e.g. achieve remission)", "two years in specialized mental health care; received at least one psychological treatment and at least three medication steps according to the national multidisciplinary guidelines on anxiety and depressive disorders"

[1] Commonly used measurement instruments, diagnoses, treatment modalities (e.g.: SSRIs, CBT) were abbreviated in this table even if the original paper did not abbreviate them.

eTable 4. Definitions for treatment resistance in anxiety disorders, per included study and per criterion.

		Psychotherapy		Psychopharmacological treatment									
Authors, year	**Minimal number of treatments**	**Unspecified[1]**	**CBT**	**Unspe-cified[1]**	**SSRIs and/ or SNRIs**	**other**	**Other biological treatments**	**Minimal length of treatment**	**Treatment response criterion**	**Duration of disorder**	**Severity of symptoms**	**Functional impair-ment**	**Comorbidity**
Aarre, 2003	x	v		v			x	x	CGI-I ≤2	x	x	x	x
Bakish et al., 1995	2	x	x		v	v	x	4 weeks	x	x	x	x	x
Bakker et al., 2005in order to address the questions (1	1	x	x		v	x	x	4 weeks	x	x	x	x	x
Baldwin & Polkinghorn, 2005	x	x	x	v			x	x	x	x	x	x	x
Bandelow et al., 2002	1	x	x	v			x	4-6 weeks	x	x	x	x	x
Bandelow, 2008	1	x	x	v			x	4-6 weeks	ΔHAM-A ≥50%	x	x	x	x
Bandelow et al., 2008	1	x	x	v			x	4-6 weeks	ΔHAM-A ≥50%	x	x	x	x
Barton et al., 2014	1	v		v			x	4 weeks	Δanxiety scale ≥50%	x	x	x	x
Brawman-Mintzer et al., 2005	1	x	x		v	v	x	4 weeks	x	x	HAM-A ≥18 (items 1&2 ≥2), Covi Anxiety Scale: moderate, CGI-S ≥4	x	x
Bystritsky, 2006	x	v		v			x	x	x	x	x	v	x
Castle et al., 2017	x	x	x	x	x	x	x	x	x	x	HAMA≥20 LSAS≥60	x	x
Chen & Tsai, 2016	1	x	x	x	x	x	x	6 months	1: No panic attacks, ≤mild avoidance, HAM-A ≤7-10, SDS ≤1 on each item, HAM-D ≤7, or 2: PDSS ≤3 and no item >1	x	1: HAM-A >7-10, HAM-D >7, or, 2: PDSS >3, or any item >1	1: SDS >1 on each item, or 2: x	1: x, or 2: HAM-D ≤7
Cosci & Fava, 2013	from 1 to over 3	from one trial (stage 1) to at least three trials including combination treatment (stage 4)					x	x	x	x	x	x	x
Cowley et al., 1997	1	x	x	v			x	x	x	x	x	x	x

		Psychotherapy		Psychopharmacological treatment									
Authors, year	**Minimal number of treatments**	**Unspecified[1]**	**CBT**	**Unspe-cified[1]**	**SSRIs and/ or SNRIs**	**other**	**Other biological treatments**	**Minimal length of treatment**	**Treatment response criterion**	**Duration of disorder**	**Severity of symptoms**	**Functional impair-ment**	**Comorbidity**
Deligiannidis & Rothschild, 2010	x		v		v	v	x	x	x	x	x	x	x
De Salas-Cansado et al., 2013	1	x	x	v			x	6 months	ΔHAM-A ≥50%, HAM-A ≤9	x	HAM-A ≥16 & CGI-S ≥3	x	x
Durham et al., 2012	"extensive treatment"	v		v			x	x	x	x	x	x	x
Gabriel, 2010	1	x	x		v	x	x	8 weeks	HAM-A <7	x	HAM-A >7 & ΔHAM-A ≥50%	x	x
Gabriel & Violato, 2011	1	x	x		v	x	x	8 weeks	HAM-A <7	x	HAM-A >10	x	x
George et al., 2008	4		v	v			x	x	ΔHAM-A ≥50%	x	HAM-A ≥20, CGI-S ≥4	x	x
Glue et al., 2017	x	x	x		x		x	x	x	x	HAMA≥20 LSAS≥60	x	x
Glue et al., 2018	x	x	x		x		x	x	x	x	HAMA≥20 LSAS≥60	x	x
Gloster et al., 2015	1		v	v			x	CBT: 20 sessions	PAS ≤18, CGI-S ≤3	x	MI ≥1.5, CGI-S ≥4	x	x
Heldt et al., 2003	1	x	x		v	x	x	4 months	No panic attacks, CGI≤2, ΔHAM-A, Δagoraphobia, Δanticipatory anxiety ≥50%	x	x	x	x
Heldt et al., 2006	1	x	x	v			x	4 months	No panic attacks, CGI≤2, ΔHAM-A, Δagoraphobia, Δanticipatory anxiety ≥50%	x	CGI-S ≥3	x	x
Hirschmann et al., 2000	3	x	x	v	v	x	x	8 weeks	x	x	ΔPSQ & CAS+PA ≥20%	x	x

		Psychotherapy		Psychopharmacological treatment									
Authors, year	**Minimal number of treatments**	**Unspecified[1]**	**CBT**	**Unspe-cified[1]**	**SSRIs and/ or SNRIs**	**other**	**Other biological treatments**	**Minimal length of treatment**	**Treatment response criterion**	**Duration of disorder**	**Severity of symptoms**	**Functional impair-ment**	**Comorbidity**
Hoge et al., 2008	1	x	x		v	v	x	8 weeks	x	x	GAD: HAM-A ≥16, CGI-S ≥4; PD: MGHAP CGI-S ≥4	x	x
Hollifield et al., 2005	2		v	v			x	8 weeks	x	x	x	x	x
Holt & Lydiard, 2007	1		v	v			x	x	x	x	x	x	x
Ipser et al., 2006	TR: 1 refr: 2	TR: x refr: v		refr: v	TR: v	x	x	x	used a variety	x	x	x	x
Katzman et al., 2008	1	x	x	v			x	8 weeks	HAM-A ≤10	x	CGI-S ≥5, HAM-A ≥20, >2 on anxious mood and tension	x	x
Kinrys et al., 2007	1	x	x	v			x	x	CGI-I ≤2	x	CGI-S ≥4	x	x
Kinrys et al., 2007	1	x	x		v	v	x	8 weeks	CGI-I ≤2	x	CGI-S ≥4	x	x
Lohoff et al., 2010r	1	x	x		v	v	x	x	x	x	HAM-A ≥16, CGI-S ≥4	x	x
Lorenz et al., 2010	1	x	x	v			x	x	x	x	x	x	x
Menza et al., 2007	x	x	x	v			x	6 weeks	ΔHAM-A ≥50%	x	HAM-A ≥14, CGI-S ≥4	x	x
Milrod et al., 2016	1		v		v	v	x	2 months (ADs), 11 weeks (CBT)	x	x	HAM-A >15	x	x
National Institute for Health and Clinical Excellence, 2011	3 steps (in stepped care algorithm)		v		v	v	x	x	x	x	"severe"	"marked"	x
Ociskova et al., 2016	x	v		v			x	x	CGI-S ≤2	x	x	x	x
Otto et al., 1999	1	x	x		v	v	x	2 months	CGI-S ≤2	x	CGI-S ≥4	x	x

		Psychotherapy		Psychopharmacological treatment									
Authors, year	**Minimal number of treatments**	**Unspecified[1]**	**CBT**	**Unspe-cified[1]**	**SSRIs and/ or SNRIs**	**other**	**Other biological treatments**	**Minimal length of treatment**	**Treatment response criterion**	**Duration of disorder**	**Severity of symptoms**	**Functional impair-ment**	**Comorbidity**
Pallanti & Quercioli, 2006	1	x	x		v	x	x	12 weeks	x	1 year	ΔLSAS <35%, CGI-I ≥3	x	x
Patterson & Van Ameringen, 2016	1	x	x		v	x	x	4-5 weeks	CGI-S ≤2	x	SAD: ΔLSAS <50%, GAD/ PD: ΔHAM-A <50%, PD: ΔPDSS <50%	x	x
Pollack et al., 1994	1	x	x		v	v	x	12 weeks	x	x	x	x	x
Pollack et al., 2006	1	x	x		v	x	x	6 weeks	ΔHAM-A ≥50%, HAM-A ≤7, CGI-S ≤2	x	CGI-S ≥4, ΔHAM-A <50%	x	x
Pollack, 2009	1	x	x		v	v	x	x	x	x	x	x	x
Rickels et al., 2012	2	x	x		v	x	x	8 weeks	ΔHAM-A ≥50%	x	HAM-A >16	x	x
Samuel et al., 2011	1	x	x	v			x	x	x	x	x	x	x
Simon et al., 2006	1	x	x		v	v	x	8 weeks	x	x	GAD: HAM-A ≥16, SAD: LSAS ≥70, PD: Panic CGI-S ≥4, All: CGI-S ≥4	x	x
Simon et al., 2009	1	x	x		v	x	x	6 weeks	CGI-S ≤2, no panic attacks last week	x	CGI-S ≥3, ≥1 panic attack last week	x	x
Snyderman et al., 2005	1	x	x	v			x	8 weeks	ΔHAM-A ≥50%, HAM-A ≤7, CGI-I ≤2	x	CGI-S ≥4, HAM-A ≥16	x	x
Solbakken & Abbass, 2015	3	v		v			x	x	OQ-45.2 ≤63, ΔOQ-45.2 ≥14, SCL-90-R ≤0.87, ΔSCL-90-R ≥0.32, IIP-64 ≤1.37, ΔIIP-64 ≥0.25	x	x	"loss of function in multiple domains"	x

		Psychotherapy		Psychopharmacological treatment									
Authors, year	**Minimal number of treatments**	**Unspecified[1]**	**CBT**	**Unspe-cified[1]**	**SSRIs and/ or SNRIs**	**other**	**Other biological treatments**	**Minimal length of treatment**	**Treatment response criterion**	**Duration of disorder**	**Severity of symptoms**	**Functional impair-ment**	**Comorbidity**
Solbakken & Abbass, 2016	3	v		v			x	x	x	x	x	"loss of function in multiple domains"	x
Starcevic, 2008	1 or 2	x	x		2 SSRIs	v	x	x	x	x	x	x	x
Stein et al., 2001	1	x	x		v	x	x	8-12 weeks	x	x	x	x	x
Stein et al., 2010	1	x	x		v	x	x	x	x	x	x	x	x
Stein, 2003	1	x	x	v			x	PD: 8 weeks, SAD: 12 weeks	x	x	x	x	x
Stein, 2004	1	x	x	v			x	8 weeks	x	x	x	x	x
Tesar & Rosenbaum, 1986	3	x	x		x	v	x	x	x	x	x	x	x
Van Ameringen et al., 2009	1	x	x		v	x	x	x	x	x	x	x	x
Worthington III et al., 2005	1	x	x		v	x	x	x	x	x	x	x	x
Yoshinaga et al., 2016	1	v			v	x	x	12 weeks	ΔLSAS ≥31%	x	LSAS ≥50, CGI-I ≥3	x	x
Zoun et al., 2016	4	v		v			x	x	x	2 years	x	x	x

[1] used when authors did not mention which specific type of intervention was mandatory in their definition of treatment resistance, for instance "psychotherapeutic treatment", or "antidepressants".
x= not included in definition, v= included in definition.

References

Aarre, T. F. (2003). Phenelzine efficacy in refractory social anxiety disorder: A case series. *Nordic Journal of Psychiatry*, *57*(4), 313–315. https://doi.org/10.1080/08039480310002110

Bakish, D., Hooper, C. L., West, D. L., Miller, C., Blanchard, A., & Bashir, F. (1995). Moclobemide and specific serotonin re-uptake inhibitor combination treatment of resistant anxiety and depressive disorders. *Human Psychopharmacology: Clinical and Experimental*, *10*(2), 105–109. https://doi.org/10.1002/hup.470100205

Bakker, A., Van Balkom, A. J. L. M., & Stein, D. J. (2005). Evidence-based pharmacotherapy of panic disorder. *International Journal of Neuropsychopharmacology*, *8*(3), 473–482. https://doi.org/10.1017/S1461145705005201

Baldwin, D. S., & Polkinghorn, C. (2005). Evidence-based pharmacotherapy of generalized anxiety disorder. In D. J. Stein, B. Lerer, S. Stahl, D. J. Stein (Ed), B. Lerer (Ed), & S. Stahl (Ed) (Eds.), *Evidence-based psychopharmacology.* (pp. 88–104). New York, NY, US: Cambridge University Press.

Bandelow, B. (2008). The medical treatment of obsessive-compulsive disorder and anxiety. *CNS Spectrums*, *13*(9, 14), 37–47.

Bandelow, B., Zohar, J., Hollander, E., Kasper, S., & Moller, H.-J. (2002). World Federation of Societies of Biological Psychiatry (WfSBP) Guidelines for the Pharmacological Treatment of Anxiety, Obsessive-Compulsive and Psttraumatic Stress Disorders. *World J Biol Psychiatry*, *3*(904758812), 171–199. https://doi.org/10.1080/15622970802465807

Bandelow, B., Zohar, J., Hollander, E., Kasper, S., Möller, H.-J., & WFSBP Task force on treatment guidelines for anxiety obsessive-compulsive post-traumatic stress disorders. (2008). World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for the pharmacological treatment of anxiety, obsessive-compulsive and post-traumatic stress disorders - First revision. *World Journal of Biological Psychiatry*, *9*(4), 248–312. https://doi.org/10.1080/15622970802465807

Barton, S., Karner, C., Salih, F., Baldwin, D. S., & Edwards, S. J. (2014). Clinical effectiveness of interventions for treatment-resistant anxiety in older people: A systematic review. *Health Technology Assessment*, *18*(50), 1–62. https://doi.org/10.3310/hta18500

Brawman-Mintzer, O., Knapp, R. G., & Nietert, P. J. (2005). Adjunctive risperidone in generalized anxiety disorder: A double-blind, placebo-controlled study. *Journal of Clinical Psychiatry*, *66*(10), 1321–1325.

Bystritsky, A. (2006). Treatment-resistant anxiety disorders. *Molecular Psychiatry*, *11*(9), 805–814. https://doi.org/10.1038/sj.mp.4001852

Castle, C., Gray, A., Neehoff, S., & Glue, P. (2017). Effect of ketamine dose on self-rated dissociation in patients with treatment refractory anxiety disorders. *Journal of Psychopharmacology*, *31*(10), 1306–1311. https://doi.org/10.1177/0269881117725685

Chen, M.-H., & Tsai, S.-J. (2016). Treatment-resistant panic disorder: clinical significance, concept and management. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *70*, 219–226. https://doi.org/10.1016/j.pnpbp.2016.02.001

Cosci, F., & Fava, G. A. (2013). Staging of mental disorders: systematic review. *Psychotherapy and Psychosomatics*, *82*(1), 20–34. https://doi.org/10.1159/000342243

Cowley, D. S., Ha, E. H., & Roy-Byrne, P. P. (1997). Determinants of pharmacologic treatment failure in panic disorder. *Journal of Clinical Psychiatry*, *58*(12), 555–561.

De Salas-Cansado, M., Álvarez, E., Olivares, J. M., Carrasco, J. L., Ferro, M. B., & Rejas, J. (2013). Modelling the cost-effectiveness of pregabalin versus usual care in daily practice in the treatment of refractory generalised anxiety disorder in Spain. *Social Psychiatry and Psychiatric Epidemiology*, *48*(6), 985–996. https://doi.org/10.1007/s00127-012-0606-6

Deligiannidis, K. M., & Rothschild, A. J. (2010). Mood and anxiety disorders. In A. J. Rothschild (Ed.), *The evidence-based guide to antipsychotic medications.* (1st ed., pp. 45–100). Arlington, VA, US: American Psychiatric Publishing, Inc.

Durham, R. C., Higgins, C., Chambers, J. A., Swan, J. S., & Dow, M. G. T. (2012). Long-term outcome of eight clinical trials of CBT for anxiety disorders: Symptom profile of sustained recovery and treatment-resistant groups. *Journal of Affective Disorders*, *136*(3), 875–881. https://doi.org/10.1016/j.jad.2011.09.017

Gabriel, A. (2010). The mixed amphetamine salt extended release (Adderall XR, Max-XR) as an adjunctive to SSRIs or SNRIs in the treatment of adult ADHD patients with comorbid partially responsive generalized anxiety: an open-label study. *ADHD Attention Deficit and Hyperactivity Disorders*, *2*(2), 87–92. https://doi.org/10.1007/s12402-010-0025-z

Gabriel, A., & Violato, C. (2011). Adjunctive atomoxetine to SSRIs or SNRIs in the treatment of adult ADHD patients with comorbid partially responsive generalized anxiety (GA): An open-label study. *ADHD Attention Deficit and Hyperactivity Disorders*, *3*(4), 319–326. https://doi.org/10.1007/s12402-011-0063-1

George, M. S., Ward, H. E., Ninan, P. T., Pollack, M., Nahas, Z., Anderson, B., … Ballenger, J. C. (2008). A pilot study of vagus nerve stimulation (VNS) for treatment-resistant anxiety disorders. *Brain Stimulation*, *1*(2), 112–121. https://doi.org/10.1016/j.brs.2008.02.001

Gloster, A. T., Sonntag, R., Hoyer, J., Meyer, A. H., Heinze, S., Ströhle, A., ... Wittchen, H.-U. (2015). Treating treatment-resistant patients with panic disorder and agoraphobia using psychotherapy: A randomized controlled switching trial. *Psychotherapy and Psychosomatics*, *84*(2), 100–109. https://doi.org/10.1159/000370162

Glue, P., Medlicott, N. J., Harland, S., Neehoff, S., Anderson-Fahey, B., Le Nedelec, M., ... McNaughton, N. (2017). Ketamine's dose-related effects on anxiety symptoms in patients with treatment refractory anxiety disorders. *Journal of Psychopharmacology*, *31*(10), 1302–1305. https://doi.org/10.1177/0269881117705089

Glue, P., Neehoff, S. M., Medlicott, N. J., Gray, A., Kibby, G., & McNaughton, N. (2018). Safety and efficacy of maintenance ketamine treatment in patients with treatment-refractory generalised anxiety and social anxiety disorders. *Journal of Psychopharmacology*, 269881118762073. https://doi.org/10.1177/0269881118762073

Heldt, E., Manfro, G. G., Kipper, L., Blaya, C., Isolan, L., & Otto, M. W. (2006). One-year follow-up of pharmacotherapy-resistant patients with panic disorder treated with cognitive-behavior therapy: Outcome and predictors of remission. *Behaviour Research and Therapy*, *44*(5), 657–665. https://doi.org/10.1016/j.brat.2005.05.003

Heldt, E., Manfro, G. G., Kipper, L., Blaya, C., Maltz, S., Isolan, L., ... Otto, M. W. (2003). Treating Medication-Resistant Panic Disorder: Predictors and Outcome of Cognitive-Behavior Therapy in a Brazilian Public Hospital. *Psychotherapy and Psychosomatics*, *72*(1), 43–48. https://doi.org/10.1159/000067188

Hirschmann, S., Dannon, P. N., Iancu, I., Dolberg, O. T., Zohar, J., & Grunhaus, L. (2000). Pindolol Augmentation in Patients With Treatment-Resistant Panic Disorder: A Double-Blind, Placebo-Controlled Trial. *Journal of Clinical Psychopharmacology*, *20*(5), 556–559. https://doi.org/10.1097/00004714-200010000-00011

Hoge, E. A., Worthington III, J. J., Kaufman, R. E., Delong, H. R., Pollack, M. H., Simon, N. M., ... Simon, N. M. (2008). Aripiprazole as augmentation treatment for refractory generalized anxiety disorder and panic disorder. *CNS Spectrums*, *13*(6), 522–527.

Hollifield, M., Thompson, P. M., Ruiz, J. E., & Uhlenhuth, E. H. (2005). Potential effectiveness and safety of olanzapine in refractory panic disorder. *Depression and Anxiety*, *21*(1), 33–40. https://doi.org/10.1002/da.20050

Holt, R. L., & Lydiard, R. B. (2007). Management of treatment-resistant panic disorder. *Psychiatry*, *4*(10), 48–59.

Ipser, J. C., Carey, P., Dhansay, Y., Fakier, N., Seedat, S., & Stein, D. J. (2006). Pharmacotherapy augmentation strategies in treatment-resistant anxiety disorders (Review). *Cochrane Database of Systematic Reviews*, (4). https://doi.org/10.1002/14651858.CD005473.pub2

Katzman, M. A., Vermani, M., Jacobs, L., Marcus, M., Kong, B., Lessard, S., ... Gendron, A. (2008). Quetiapine as an adjunctive pharmacotherapy for the treatment of non-remitting generalized anxiety disorder: A flexible-dose, open-label pilot trial. *Journal of Anxiety Disorders*, *22*(8), 1480–1486. https://doi.org/10.1016/j.janxdis.2008.03.002

Kinrys, G., Vasconcelos E Sa, D., & Nery, F. (2007). Adjunctive zonisamide for treatment refractory anxiety. *International Journal of Clinical Practice*, *61*(6), 1050–1053. https://doi.org/10.1111/j.1742-1241.2007.01365.x

Kinrys, G., Worthington, J. J., Wygant, L., Nery, F., Reese, H., & Pollack, M. H. (2007). Levetiracetam as adjunctive therapy for refractory anxiety disorders. *Journal of Clinical Psychiatry*, *68*(7), 1010–1013. Retrieved from http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L47340319

Lohoff, F. W., Etemad, B., Mandos, L. A., Gallop, R., & Rickels, K. (2010). Ziprasidone Treatment of Refractory Generalized Anxiety Disorder. *Journal of Clinical Psychopharmacology*, *30*(2), 185–189. https://doi.org/10.1097/JCP.0b013e3181d21951

Lorenz, R. A., Jackson, C. W., & Saitz, M. (2010). Adjunctive use of atypical antipsychotics for treatment-resistant generalized anxiety disorder. *Pharmacotherapy*, *30*(9), 942–951. https://doi.org/10.1592/phco.30.9.942

Menza, M. A., Dobkin, R. D., & Marin, H. (2007). An open-label trial of aripiprazole augmentation for treatment-resistant generalized anxiety disorder [3]. *J Clin Psychopharmacol*, *22*(2), 207–210. https://doi.org/10.1097/01.jcp.0000248620.34541.bc

Milrod, B., Altemus, M., Gross, C., Busch, F., Silver, G., Christos, P., ... Schneier, F. (2016). Adult separation anxiety in treatment nonresponders with anxiety disorders: Delineation of the syndrome and exploration of attachment-based psychotherapy and biomarkers. *Comprehensive Psychiatry*, *66*, 139–145. https://doi.org/10.1016/j.comppsych.2016.01.004

National Institute for Health and Clinical Excellence. (2011). *Generalised anxiety disorder and panic disorder in adults: management*. *CG113*. Manchester, United Kingdom.

Ociskova, M., Prasko, J., Latalova, K., Kamaradova, D., & Grambal, A. (2016). Psychological factors and treatment effectiveness in resistant anxiety disorders in highly comorbid inpatients. *Neuropsychiatric Disease and Treatment*, *12*, 1539–1551. https://doi.org/10.2147/NDT.S104301

Otto, M. W., Pollack, M. H., Penava, S. J., & Zucker, B. G. (1999). Group cognitive-behavior therapy for patients failing to respond to pharmacology for panic disorder: A clinical case series. *Behaviour Research and Therapy*, *37*(8), 763–770. https://doi.org/10.1016/S0005-7967(98)00176-4

Pallanti, S., & Quercioli, L. (2006). Resistant social anxiety disorder response to Escitalopram. *Clinical Practice and Epidemiology in Mental Health*, *2*(35). https://doi.org/10.1186/1745-0179-2-35

Patterson, B., & Van Ameringen, M. (2016). AUGMENTATION STRATEGIES FOR TREATMENT-RESISTANT ANXIETY DISORDERS: A SYSTEMATIC REVIEW AND META-ANALYSIS. *Depression and Anxiety*, *33*(8), 728–736. https://doi.org/10.1002/da.22525

Pollack, M. H. (2009). Refractory Generalized Anxiety Disorder. *The Journal of Clinical Psychiatry*, *70*(suppl 2), 32–38. https://doi.org/10.4088/JCP.s.7002.06

Pollack, M. H., Otto, M. W., Kaspi, S. P., Hammerness, P. G., & Rosenbaum, J. F. (1994). Cognitive behavior therapy for treatment-refractory panic disorder. *Journal of Clinical Psychiatry*, *55*(5), 200–205. Retrieved from http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L24184829

Pollack, M. H., Simon, N. M., Zalta, A. K., Worthington, J. J., Hoge, E. A., Mick, E., … Oppenheimer, J. (2006). Olanzapine augmentation of fluoxetine for refractory generalized anxiety disorder: A placebo controlled study. *Biological Psychiatry*, *59*(3), 211–215. https://doi.org/10.1016/j.biopsych.2005.07.005

Rickels, K., Shiovitz, T. M., Ramey, T. S., Weaver, J. J., Knapp, L. E., & Miceli, J. J. (2012). Adjunctive therapy with pregabalin in generalized anxiety disorder patients with partial response to SSRI or SNRI treatment. *International Clinical Psychopharmacology*, *27*(3), 142–150. https://doi.org/10.1097/YIC.0b013e328350b133

Samuel, M., Zimovetz, E. A., Gabriel, Z., & Beard, S. M. (2011). Efficacy and safety of treatments for refractory generalized anxiety disorder: A systematic review. *International Clinical Psychopharmacology*, *26*(2), 63–68. https://doi.org/10.1097/YIC.0b013e328341bb4a

Simon, N. M., Hoge, E. A., Fischmann, D., Worthington III, J. J., Christian, K. M., Kinrys, G., & Pollack, M. H. (2006). An open-label trial of risperidone augmentation for refractory anxiety disorders. *Journal of Clinical Psychiatry*, *67*(3), 381–385. Retrieved from http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L43556561

Simon, N. M., Otto, M. W., Worthington, J. J., Hoge, E. A., Thompson, E. H., LeBeau, R. T., … Pollack, M. H. (2009). Next-Step Strategies for Panic Disorder Refractory to Initial Pharmacotherapy. *The Journal of Clinical Psychiatry*, *70*(11), 1563–1570. https://doi.org/10.4088/JCP.08m04485blu

Snyderman, S. H., Rynn, M. A., Rickels, K., Snydermae, S. H., Rynn, M. A., Rickels, K., … Rickels, K. (2005). Open-label Pilot Study of Ziprasidone for Refractory Generalized Anxiety Disorder. *Journal of Clinical Psychopharmacology*, *25*(5), 497–499. https://doi.org/10.1097/01.jcp.0000177853.15910.de

Solbakken, O. A., & Abbass, A. (2015). Intensive short-term dynamic residential treatment program for patients with treatment-resistant disorders. *Journal of Affective Disorders*, *181*, 67–77. https://doi.org/10.1016/j.jad.2015.04.003

Solbakken, O. A., & Abbass, A. (2016). Symptom- and personality disorder changes in intensive short-term dynamic residential treatment for treatment-resistant anxiety and depressive disorders. *Acta Neuropsychiatrica*, *28*(05), 257–271. https://doi.org/10.1017/neu.2016.5

Starcevic, V. (2008). Treatment of panic disorder: Recent developments and current status. *Expert Review of Neurotherapeutics*, *8*(8), 1219–1232. https://doi.org/10.1586/14737175.8.8.1219

Stein, D. J. (2003). Algorithm for the pharmacotherapy of anxiety disorders. *Current Psychiatry Reports*, *5*(4), 282–288. Retrieved from http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L38899220

Stein, D. J. (2004). Algorithms for primary care: An evidence-based approach to the pharmacotherapy of depression and anxiety disorders. *Primary Psychiatry*, *11*(6), 55–78. Retrieved from http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L38879037

Stein, D. J., Baldwin, D. S., Bandelow, B., Blanco, C., Fontenelle, L. F., Lee, S., ... Van Ameringen, M. (2010). A 2010 evidence-based algorithm for the pharmacotherapy of social anxiety disorder. *Current Psychiatry Reports*, *12*(5), 471–477. https://doi.org/10.1007/s11920-010-0140-8

Stein, D. J., Kasper, S., Matsunaga, H., Osser, D. N., Stein, M. B., Van Ameringen, M., ... Zhang, M. (2001). Pharmacotherapy of social anxiety disorder: an algorithm for primary care - 2001. *Primary Care Psychiatry*, *7*(3), 107–110. https://doi.org/10.1185/135525701317094368

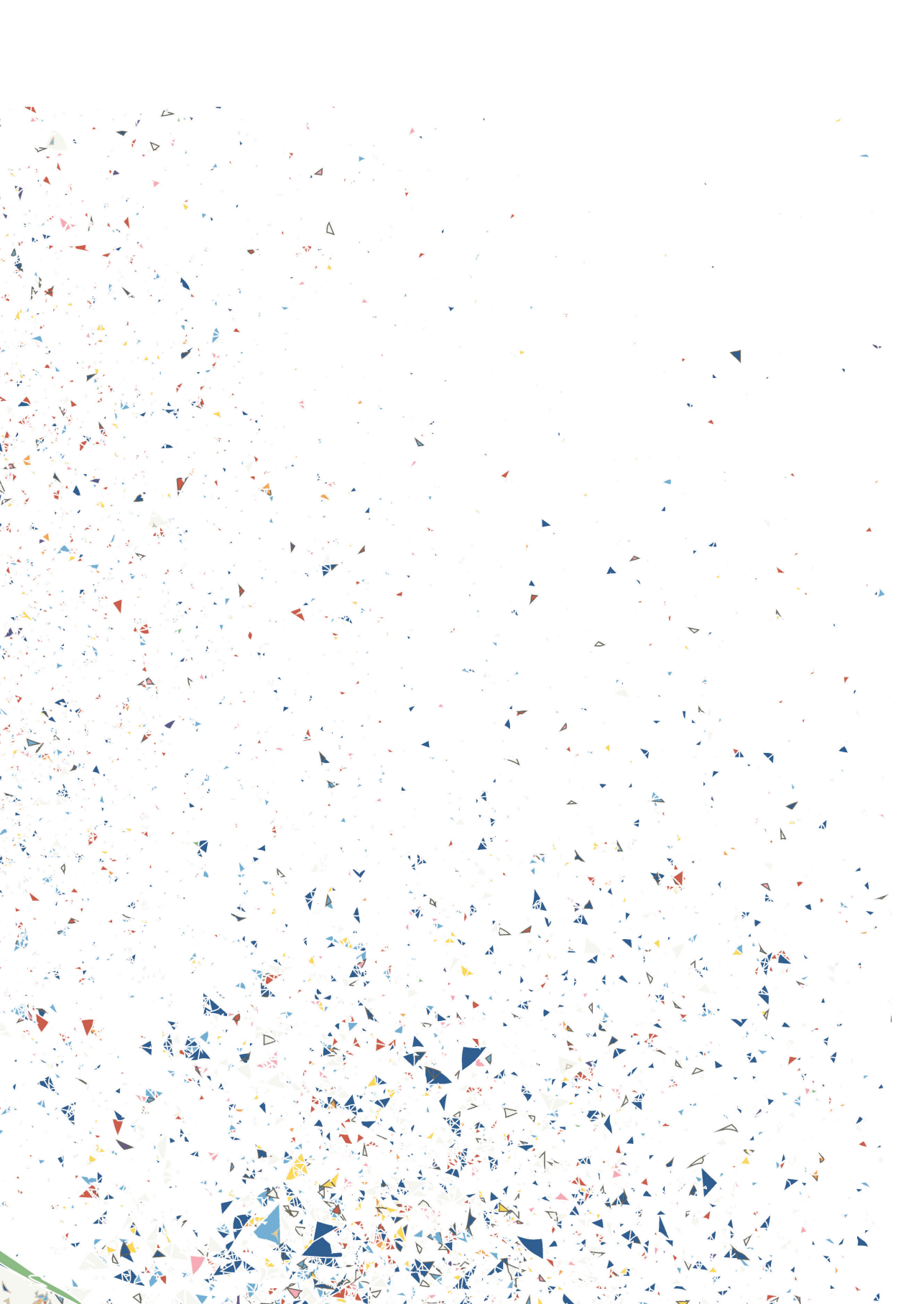
Tesar, G. E., & Rosenbaum, J. F. (1986). Successful use of clonazepam in patients with treatment-resistant panic disorder. *Journal of Nervous and Mental Disease*, *174*(8), 477–482. Retrieved from http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L16040140

Van Ameringen, M., Mancini, C., & Patterson, B. (2009). Pharmacotherapy for social anxiety disorder and specific phobia. In M. M. Antony, M. B. Stein, M. M. Antony (Ed), & M. B. Stein (Ed) (Eds.), *Oxford handbook of anxiety and related disorders.* (pp. 321–333). New York, NY, US: Oxford University Press. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2008-13193-024&site=ehost-live

Worthington III, J. J., Kinrys, G., Wygant, L. E., & Pollack, M. H. (2005). Aripiprazole as an augmentor of selective serotonin reuptake inhibitors in depression and anxiety disorder patients. *International Clinical Psychopharmacology*, *20*(1), 9–11. https://doi.org/10.1097/00004850-200501000-00002

Yoshinaga, N., Matsuki, S., Niitsu, T., Sato, Y., Tanaka, M., Ibuki, H., ... Shimizu, E. (2016). Cognitive behavioral therapy for patients with social anxiety disorder who remain symptomatic following antidepressant treatment: A randomized, assessor-blinded, controlled trial. *Psychotherapy and Psychosomatics*, *85*(4), 208–217. https://doi.org/10.1159/000444221

Zoun, M. H. H., Koekkoek, B., Sinnema, H., Muntingh, A. D. T., van Balkom, A. J. L. M. A. J. L. M., Schene, A. H., ... Spijker, J. (2016). Effectiveness and cost-effectiveness of a self-management training for patients with chronic and treatment resistant anxiety or depressive disorders: design of a multicenter randomized controlled trial. *BMC Psychiatry*, *16*(1). https://doi.org/10.1186/s12888-016-0927-1

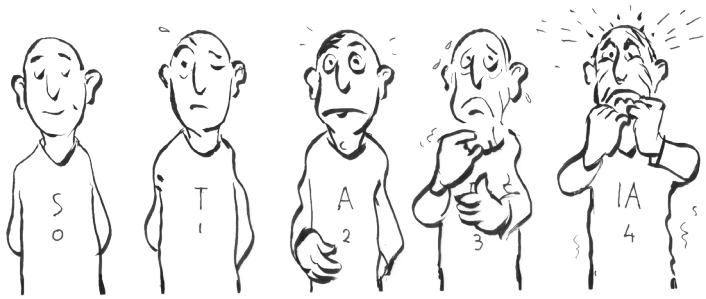


Chapter 5.

A clinical staging approach to improving diagnostics in anxiety disorders: is it the way to go?

Australian and New Zealand Journal of Psychiatry (2020).

Wicher A. Bokma
Neeltje M. Batelaan
Adriaan W. Hoogendoorn
Brenda W.J.H. Penninx
Anton J.L.M. van Balkom



Abstract

Background

Clinical staging is a paradigm in which stages of disease progression are identified; these, in turn, have prognostic value. A staging model that enables the prediction of long-term course in anxiety disorders is currently unavailable but much needed as course trajectories are highly heterogenic. This study therefore tailored a heuristic staging model to anxiety disorders and assessed its validity.

Methods

A clinical staging model was tailored to anxiety disorders, distinguishing nine stages of disease progression varying from subclinical stages (0, 1A, 1B) to clinical stages (2A-4B). At-risk subjects and subjects with anxiety disorders (n = 2,352) from the longitudinal Netherlands Study of Depression and Anxiety (NESDA) were assigned to these nine stages. The model's validity was assessed by comparing baseline (construct validity) and two-year, four-year and six-year follow-up (predictive validity) differences in anxiety severity measures across stages. Differences in depression severity and disability were assessed as secondary outcome measures.

Results

Results showed that the anxiety disorder staging model has construct and predictive validity. At baseline, differences in anxiety severity, social avoidance behaviors, agoraphobic avoidance behaviors, worrying, depressive symptoms and levels of disability existed across all stages (all p-values <0.001). Over time, these differences between stages remained present until the six-year follow-up. Differences across stages followed a linear trend in all analyses: higher stages were characterized by the worst outcomes. Regarding the stages, subjects with psychiatric comorbidity (stages 2B, 3B, 4B) showed a deteriorated course compared with those without comorbidity (stages 2A, 3A, 4A).

Conclusions

A clinical staging tool would be useful in clinical practice to predict disease course in anxiety disorders.

Introduction

Anxiety disorders are characterized by highly heterogenic course trajectories. The longitudinal course of anxiety disorders is characterized as "chronic" in almost 60% of the cases after two years [1] and in 20% to 60% of the cases after 12 years.[2] In contrast, another subset of anxiety disorder patients experience a mild course with moderate symptom severity and lower subjective need for care.[1,3] Information on prognosis is essential to tailor treatment to individual needs. However, such information is not provided by a DSM classification. Any tool that contributes to course prediction in anxiety disorders would be of great clinical relevance.[4–6]

In clinical staging, stages are distinguished that reflect increasing levels of disease progression. Disorders of individual patients are assigned to a certain stage according to their risk profile. One of the most widely recognized clinical staging models for psychiatric disorders is the heuristic model developed by McGorry and colleagues.[7,8] This model covers the spectrum from asymptomatic, at-risk subjects (stage 0) to severe, chronic illness (stage 4). This model showed some promise in predicting the longitudinal course in adolescents seeking help for a variety of psychiatric disorders and in adults with major depressive disorder.[8,9] A study in young adults who presented with social anxiety showed clinical applicability.[10] Until now, no staging model has been studied empirically in adult patients with anxiety disorders.

In this study, we tailored McGorry's staging model to anxiety disorders and examined its validity in a heterogeneous anxiety disorder sample. For a staging model to be valid, two assumptions must be met: (i) with each successive stage, probabilities of unfavorable disease markers should increase; (ii) with each successive stage, longitudinal course should worsen.[4] The first assumption was tested cross-sectionally (construct validity), while the second assumption was tested with longitudinal data (predictive validity). In these comparisons, we distinguished two sets of validators: (a) anxiety severity measures and (b) general psychopathology measures. These sets of validators were chosen as the symptoms for anxiety disorders often overlap with those in other common mental disorders. We hypothesized that distributions of construct and predictive validators in our model would show linearity across stages with increasing baseline and follow-up severity in higher stages. This would implicate that the model is valid. If a staging model in anxiety disorders appears to be valid, treatment

guidelines could be improved by differentiating treatment according to the level of disease progression (e.g. Andrews et al., 2018; Bandelow et al., 2008; National Institute for Health and Clinical Excellence, 2013).[11–13]

Methods

Sample

Data were derived from the Netherlands Study of Depression and Anxiety (NESDA), a naturalistic longitudinal cohort study, using a sample (n = 2,981) that is representative of the various developmental stages of depression and anxiety. An extensive description of the study design for NESDA is provided elsewhere.[14]

For the purpose of this study, we included subjects with anxiety disorders (n = 1,305). Subjects without anxiety disorders were included if they had risk factors for development of anxiety disorders (n = 1,115: see below for definition). Healthy controls without risk factors were excluded (n = 156). Subjects with a diagnosis of major depressive disorder (MDD) were excluded if they did not have an anxiety disorder (n = 396). Subjects with missing data (n = 77) were also excluded. The total sample size for this study was n = 2,352. Of these, 2,042 (86.8%) were reassessed after two years (T2), 1,895 (80.6%) after four years (T4) and 1,772 (75.3%) after six year of follow-up (T6).

Measurements

Clinical staging model

The present study was based on a staging model developed by McGorry and colleagues, which we tailored to anxiety disorders.[7,8] The subjects were assigned to a certain stage in the staging model, based on their baseline measurements. The different stages included stage 0 (asymptomatic), 1A (nonspecific symptoms), 1B (attenuated syndromes), 2A (discrete disorder), 2B (discrete disorder with comorbidity), 3A (intermittent symptoms), 3B (intermittent symptoms with comorbidity), 4A (chronic symptoms), and 4B (chronic symptoms with comorbidity).

Stages 0, 1A and 1B were labelled "subclinical" stages and were assigned in subjects without DSM-IV anxiety disorders but who did have at least one risk factor for developing an anxiety disorder. Three risk factors for developing an anxiety disorder were derived from the literature: lifetime history of

Table 1. Adaptation of the staging model.

Stages Hickie et al.:	Definition:	Adapted stages for anxiety disorders:	Assignment criteria:
Stage 0	'asymptomatic subjects', at risk, no current anxiety symptoms	**Stage 0**	Presence of anxiety disorder risk factor ¤ Absence of anxiety disorder Low anxiety severity §
Stage 1A	'help-seeking', nonspecific symptoms of mild to moderate anxiety.	**Stage 1A**	Presence of anxiety disorder risk factors ¤ Absence of anxiety disorder Mild to moderate anxiety severity §
Stage 1B	'attenuated syndromes', specific symptoms of for instance severe anxiety or moderate depression.	**Stage 1B**	Criteria for stage 1A, but with moderate to severe anxiety severity §
Stage 2	'discrete disorders'.	**Stage 2A** *no comorbid disorders*	Presence of at least one anxiety disorder No chronic duration (<30% of the time presence of symptoms during 4 years before baseline)
		Stage 2B *comorbid disorder present*	Criteria for stage 2A, with comorbid current MDD, Dysthymia or Alcohol Dependency
Stage 3	'intermittent or persistent disorder'	**Stage 3A** *no comorbid disorders*	Presence of at least one anxiety disorder intermittent symptoms (30-80% of the time presence of symptoms during 4 years before baseline)
		Stage 3B *comorbid disorder present*	Criteria for stage 3A, with comorbid current MDD, Dysthymia or Alcohol Dependency
Stage 4	'severe, chronic and unremitting illness'	**Stage 4A** *no comorbid disorders*	Presence of at least one anxiety disorder Chronic course (>80% of time presence of symptoms during the 4 years prior to baseline measurement)
		Stage 4B *comorbid disorder present*	Criteria for stage 4A, with comorbid current MDD, Dysthymia or Alcohol Dependency

This model was derived from I.B. Hickie, E.M. Scott, D.F. Hermens, et al. Applying clinical staging to young people who present for mental health care. Early Interv. Psychiatry. 7 (2013) 31–43.
¤ Anxiety disorder risk factors were either a positive family history for psychiatric disorders (family tree method) (Fyer & Weissman 1999)1998: Am. J. Med. Genet. (Neuropsychiatr. Genet., traumatic life events in youth (Childhood Trauma Inventory) (Graaf *et al.* 2002), or lifetime diagnosis of anxiety disorder, currently in remission (CIDI) (World Health Organization 1998).
BAI = Beck's Anxiety Inventory, PSWQ = Penn State Worry Questionnaire, FQ (So) = Fear Questionnaire, social phobia subscale, FQ (Ag) = Fear Questionnaire, agoraphobia subscale, MDD = Major Depressive Disorder.
§ Cut-off values anxiety severity, low: BAI <10 and PSWQ <24 and FQ (Ag) <15 and FQ (So) <12; mild to moderate: BAI 10<30, or PSWQ 24<39, or FQ (Ag) 15<19, or FQ (So) 12<18; moderate to severe: BAI ≥30, or PSWQ ≥39, or FQ (Ag) ≥19, or FQ (So) ≥18.
All diagnoses were according to DSM-IV classifications and with 6 month recency.

anxiety disorders,[15,16] exposure to childhood trauma,[15,17,18] and family history for psychiatric disorders.[15,19] Stage 0 was assigned to subjects with low symptom severity, stage 1A to subjects with mild to moderate symptom severity, and stage 1B to subjects with moderate to severe symptom severity.

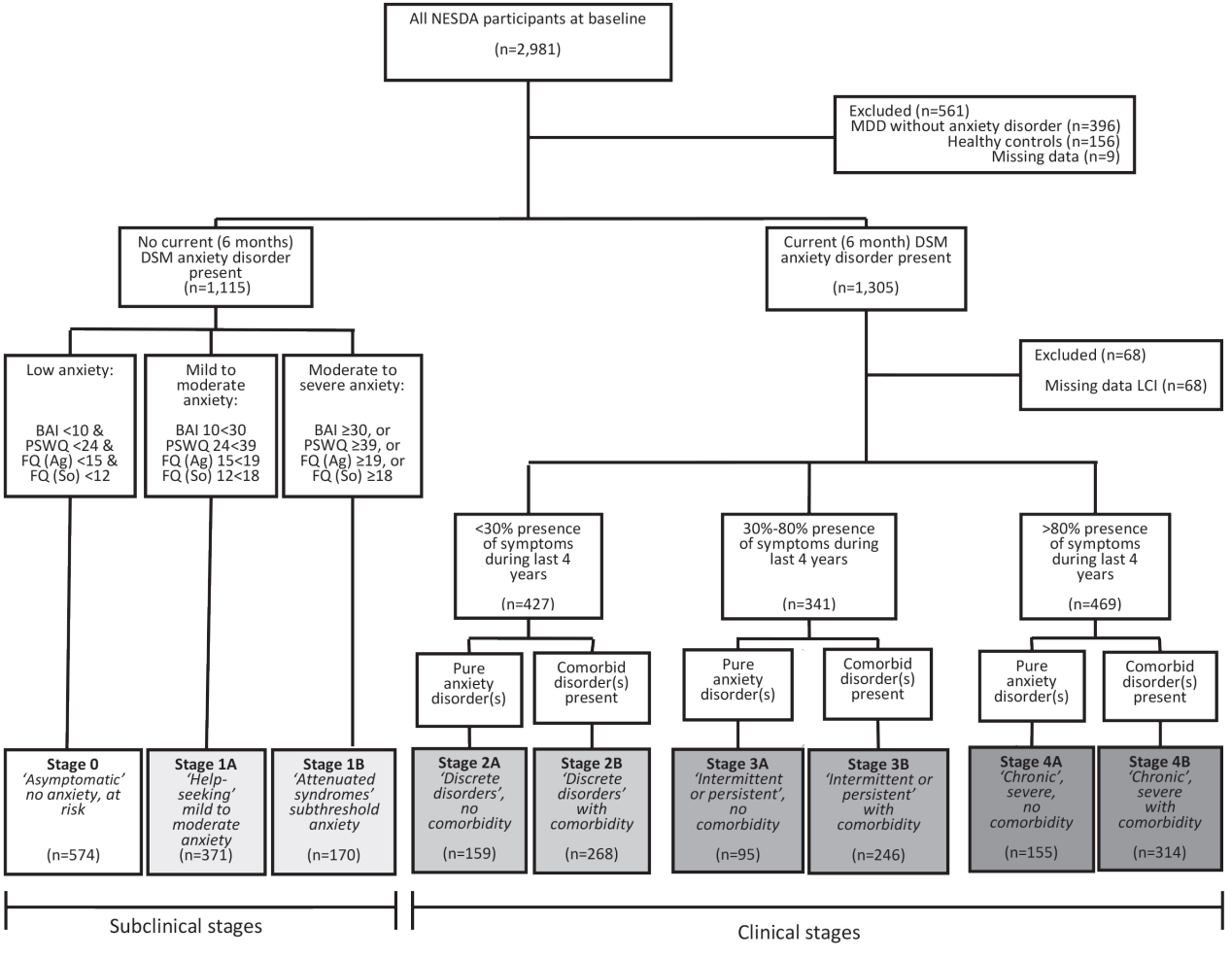


Figure 1. Flowchart for inclusion and assignment to stages.

Stages 2 through 4 were labeled the "clinical" stages. We assigned subjects to stages 2, 3 or 4 when they had any current DSM-IV anxiety disorder: either panic disorder, panic disorder with agoraphobia, agoraphobia, generalized anxiety disorder or social anxiety disorder. Subjects with a non-chronic duration prior to baseline measurements were assigned to stage 2, subjects with an intermittent duration prior to baseline measurements to stage 3, and subjects with a chronic duration prior to baseline measurements to stage 4. We adapted McGorry's model to account for the presence of psychiatric comorbidity (e.g. MDD, alcohol dependency) for two reasons. First, anxiety often co-occurs with these comorbidities,[20] and second, these comorbidities strongly predict a worse course in anxiety disorders.[1,2] Subjects with comorbid MDD, dysthymia or alcohol dependency were assigned to 'B' substages. See Table 1 for explicit assignment criteria and Figure 1 for a flowchart. See the supplementary materials for full information on the measurement instruments as well as the rationale behind the cut-off values used in stage assignment.

Instruments used for stage assignment

The World Health Organization (WHO) Composite International Diagnostic Interview (CIDI, version 2.1) was used at baseline to assess lifetime history of anxiety disorders, family history for psychiatric disorders, presence of DSM-IV panic disorder, panic disorder with agoraphobia, agoraphobia, generalized anxiety disorder, social anxiety disorder, MDD, dysthymia and alcohol dependency (six-month recency) and age of onset for these disorders. The CIDI is a structured interview with good reliability and validity.[21]

Duration of anxiety and avoidance symptoms during the last four years prior to baseline were assessed with the Life Chart Interview (LCI), a structured retrospective interview using a calendar approach.[22] The LCI has adequate reliability and validity.[23] We calculated proportional scores reflecting the duration of anxiety or avoidance symptoms: "not chronic" (<30% of months in previous four years), "intermittent" (30-80% of months in previous four years) and "chronic" (>80% of months in previous four years).

Validators

Two sets of outcome measures were used to assess construct and predictive validity: (i) anxiety severity measures (main outcome measure), and (ii) general psychopathology measures (depression severity and disability: secondary outcome measure). The latter were included since longitudinal

anxiety course is known to show high levels of comorbidity with depression and because remission of symptoms does not always indicate that the subject has recovered from the disability.[20,24-27]

Construct validation

Baseline measures were used to assess construct validity. Severity of anxiety was measured with the Beck Anxiety Inventory (BAI), a 21-item self-report questionnaire.[28] Severity of avoidance behaviors was measured with the Fear Questionnaire (FQ), a 15-item self-report questionnaire.[29] For the purpose of this study, two of its subscales were used: agoraphobic avoidance (FQ Ag) and social avoidance (FQ So). Severity of pathological worrying was assessed using the 11-item self-report version of the Penn State Worry Questionnaire (PSWQ).[30] Depressive symptoms were measured with the Inventory of Depressive Symptomatology-SR (IDS), a self-report questionnaire on severity of depression.[31] Levels of disability were measured with the WHO Disability Assessment Schedule (WHO DAS II), a 36-item self-report questionnaire measuring levels of disability.[32]

Predictive validation

At two-year, four-year and six-year follow-up, all these measurements were repeated to assess predictive validity. Furthermore, at these timepoints, presence of any anxiety disorder and presence of any psychiatric disorder (either anxiety disorder, MDD, dysthymia or alcohol dependency) were used to assess predictive validity.

Statistical analyses

Construct validity analyses

After assigning subjects to stages, baseline clinical characteristics and construct validators in our sample were compared using Pearson chi-squared statistics for dichotomous variables and one-way ANOVAs for continuous variables. Additionally, Mantel Haenszel's trend analyses and (Wilcoxon-type of) nonparametric tests for trend across ordered groups [33] were performed to investigate whether an increase across stages could be demonstrated while taking the ordinal distributions of the staging model into account. We applied Bonferroni correction (α(corrected) = α/k hypotheses) for the ten construct-validator statistical tests. This yielded an α(corrected) of 0.05/10 = 0.005. To ensure that attrition from the cohort did not lead to a power problem, the number of events per variable (EPV) at two-year, four-year and six-year timepoints were calculated for the main outcome measure.

Predictive validity analyses

Binary measures (e.g. presence of anxiety disorder) were linked to the staging model by fitting generalized estimating equations (GEE) models with an exchangeable correlation structure to longitudinal data (T2, T4 and T6), estimating effects for stages (categorical), time-points (categorical) and for all stage*time-point interactions. We adjusted these for baseline age because age was correlated with the pre-baseline duration of anxiety: younger subjects had lower duration of anxiety. From these models, odds-ratios (ORs) for presence of diagnoses at successive time-points for different stages were derived using the combined subclinical stages 0-1B at two-year follow-up as reference. For continuous measures (e.g. anxiety severity) linear mixed models were used. While GEE and mixed model analyses are quite similar, mixed models perform better in linear analyses in cases of incomplete data.[34] In these models, fixed effects for stages (categorical), time-points (categorical) and for all stage*time-point interactions were estimated while adjusting for age. These models included a random intercept. From these models, expected severity scores at successive time-points for each stage were derived.

Sensitivity analysis

Inclusion of any predictors used in stage assignment would only be justified when each showed predictive power. To check this, a multivariable logistic regression model that incorporated all stage-assignment variables independently was made to predict two-year presence of anxiety disorders. We hypothesized that all variables that were used in stage assignment were significant predictors of longitudinal course when adjusted for other predictors. See the supplement for the full rationale on this sensitivity analysis.

Results

Sample characteristics

Table 2 shows baseline demographics and clinical characteristics, stratified per stage. Gender, age, and education years were unevenly distributed across stages. The number of years of education attained was lower in higher stages ($p<0.001$). For most clinical characteristics, a pattern indicating a higher severity in higher stages was present. The exceptions were age of onset and presence of panic disorder without agoraphobia, for which negative linear associations existed (see Table 2). MDD was the

Table 2. Distribution of baseline characteristics and construct validators across stages.

	Subclinical stages			Clinical stages								
Stage	**0**	**1A**	**1B**	**2A**	**2B**	**3A**	**3B**	**4A**	**4B**	*Between groups*		
n (total =2352)	574	371	170	159	268	95	246	155	314	*Linear nonparametric trends, p*		
Demographic characteristics										*Subclinical (0-1B)*	*Clinical (2A-4B)*	*Overall (0-4B)*
Gender, % female	61.3	72.8	70.6	68.6	69.8	71.6	67.1	72.3	63.4	**<0.001**	0.190	0.314
Age (years), mean (SD)	42.6 (14.3)	43.2 (13.7)	41.1 (13.3)	38.1 (12.9)	39.2 (12.4)	40.2 (13.0)	41.3 (11.9)	44.0 (11.9)	43.8 (11.7)	0.455	**<0.001**	0.811
Education (years), mean (SD)	12.8 (3.1)	12.6 (3.3)	12.2 (3.1)	12.1 (3.0)	11.97 (3.2)	12.6 (3.5)	11.7 (3.3)	11.8 (3.3)	10.8 (3.2)	**0.047**	**<0.001**	**<0.001**
Clinical characteristics										*Subclinical (0-1B)*	*Clinical (2A-4B)*	*Overall (0-4B)*
Family history of psychiatric disorders, %	94.3	94.6	92.9	86.8	91.4	88.4	91.5	87.1	94.3	0.749	0.060	0.050
Presence of childhood trauma, %	33.4	48.5	54.1	41.5	61.6	53.7	61.0	50.3	66.6	**<0.001**	**0.001**	**<0.001**
History of anxiety disorder, %	21.8	38.0	45.9	100§	100§	100§	100§	100§	100§	**<0.001**	*na*	*na*
Age of onset anxiety disorder, mean (SD)	*na*	*na*	*na*	23.9 (13.1)	22.6 (12.9)	20.5 (12.0)	19.9 (12.2)	21.0 (13.3)	19.2 (12.4)	*na*	**<0.001***	*na*
Social phobia, %	0§	0§	0§	42.1	44.8	48.4	54.5	46.5	63.1	na	**<0.001**	*na*
Generalized anxiety disorder, %	0§	0§	0§	18.2	32.8	26.3	47.2	22.6	47.5	*na*	**0.002**	*na*
Panic disorder without agoraphobia, %	0§	0§	0§	23.9	26.1	17.9	19.9	12.3	13.7	*na*	**<0.001***	*na*
Panic disorder with agoraphobia, %	0§	0§	0§	28.3	28.7	35.8	33.3	34.2	40.1	*na*	**<0.001**	na
Agoraphobia without panic disorder, %	0§	0§	0§	11.3	11.9	10.5	11.0	27.7	15.9	*na*	**0.005**	*na*
Major depressive disorder, %	0§	0§	0§	0§	85.1	0§	89.4	0§	82.2	*na*	0.280	*na*
Dysthymia, %	0§	0§	0§	0§	11.2	0§	29.3	0§	38.9	*na*	**<0.001**	*na*
Alcohol dependency, %	7.84	11.6	11.2	0§	29.1	0§	30.5	0§	33.1	0.053	0.293	*na*
Current treatment (antidepressants or psychotherapy), %	4.20	12.1	20.6	38.4	57.8	37.9	57.7	38.1	59.9	**<0.001**	**0.017**	**<0.001**

Construct validators										*Subclinical (0-1B)*	*Clinical (2A-4B)*	*Overall (0-4B)*
Anxiety validators												
Anxiety severity (BAI), mean (SD)	2.82 (3.31)§	7.83 (5.74)§	11.7 (7.59)§	13.7 (7.88)	18.7 (10.9)	15.9 (9.24)	20.8 (10.0)	16.2 (11.1)	22.4 (11.3)	*na*	**<0.001**	*na*
Social avoidance (FQ), mean (SD)	4.65 (4.28)§	7.13 (4.84)§	15.6 (7.69)§	11.6 (8.31)	14.3 (8.57)	14.5 (9.45)	17.9 (8.33)	14.7 (9.31)	19.9 (9.32)	*na*	**<0.001**	*na*
Agoraphobic avoidance (FQ), mean (SD)	1.40 (2.73)§	2.67 (3.71)§	7.51 (7.61)§	7.24 (6.65)	9.16 (8.87)	9.31 (8.64)	11.7 (9.43)	12.3 (10.3)	15.4 (11.0)	*na*	**<0.001**	*na*
Worrying severity (PSWQ), mean (SD)	14.5 (6.39)§	27.9 (7.21)§	35.5 (13.5)§	28.7 (13.3)	30.4 (17.7)	32.7 (14.8)	32.7 (17.9)	30.1 (14.7)	33.5 (18.0)	*na*	**<0.001**	*na*
General psychopathology validators												
Depression severity (IDS), mean (SD)	7.22 (5.84)	14.3 (7.78)	21.3 (9.30)	19.6 (9.24)	29.7 (12.5)	23.8 (9.47)	34.5 (12.9)	22.7 (10.5)	35.2 (12.8)	**<0.001**	**<0.001**	**<0.001**
Disability (WHO-DAS II), mean (SD)	9.66 (11.6)	19.6 (15.7)	31.3 (18.1)	26.6 (19.7)	43.9 (20.6)	33.3 (20.1)	50.4 (22.0)	33.9 (22.7)	50.8 (24.3)	**<0.001**	**<0.001**	**<0.001**

§: cells with § were excluded from statistical tests as values in these cells were a result of assignment criteria and will therefore by design deviate from other stages. Linear trends across the stages were assessed with Cuzick's nonparametric trend statistic, p-values are given for comparisons across the subclinical stages (0, 1A, 1B), the clinical stages (2A, 2B, 3A, 3B, 4A, 4B) and across all stages. na= not available: negative linear association. BAI= Beck Anxiety Inventory, FQ= Fear Questionnaire, PSWQ= Penn State Worry Questionnaire, IDS= Inventory of Depressive Symptoms, WHO-DAS= WHO-Disability Assessment Schedule. Boldface indicates statistically significant p-values, with* α*=0.05 for demographics and clinical variables, and* $\alpha_{(corrected)}$*=0.005 for construct validators.*

most prevalent comorbidity (82.2-89.4%), followed by alcohol dependency (29.1-33.1%) and dysthymia (11.2-38.9%). Presence of MDD and alcohol dependency was evenly distributed across stages 2B, 3B and 4B, whereas dysthymia was more prevalent in higher stages (linear trend $X^2(1) = 55.3$, $p<0.001$). Furthermore, subjects in higher stages had greater chances of receiving current psychiatric treatments (Table 2). Percentages of subjects receiving treatment were highest in the comorbidity stages 2B, 3B, 4B (57.7-59.9%).

Construct validation

We tested the assumption that with each successive stage at baseline, probabilities of unfavorable disease markers would increase. In all anxiety measures, this dose-response pattern was found; namely, increasing anxiety severity in higher stages. After Bonferroni correction, linear trends were significant for all construct validators. The only exception was stage 4A, which was associated with levels of severity comparable to subjects in stage 3A; for example, mean anxiety severity in stage 3A = 15.9 (SD = 9.24), in stage 4A = 16.2 (SD = 11.1); mean social avoidance in stage 3A = 14.5 (SD = 9.45), in stage 4A = 14.7 (SD = 9.31). For the general psychopathology measures, the patterns were somewhat different. There was a gradual increase in levels of general psychopathology until stage 3, after which they remained constant. As expected, comorbidity stages (2B, 3B and 4B) all showed substantially higher baseline severity scores than non-comorbid stages (2A, 3A and 4A). See Table 2 for means and standard deviations of these measures across stages at baseline.

Predictive validation

The second assumption we tested was that with each successive stage, longitudinal course would worsen. At two-year follow-up, proportions of subjects with an anxiety disorder ranged from 2.7% (stage 0) to 68.0% (stage 4B). At four-year follow-up, anxiety disorders were present in 3.0% (stage 0) to 59.0% (stage 4B), at six-year follow-up in 3.1% (stage 0) to 55.1% (stage 4B). These were incident disorders, recurrent disorders and persistent disorders. This amounts to 78.4 (two-year) to 47.5 (six-year) events per variable (EPV). For all stages, proportions of anxiety disorders were lowest at six-year follow-up, followed by four-year and two-year follow-up. Figure 2 shows GEE derived age-adjusted ORs for presence of follow-up anxiety disorders at different time-points, using the combined subclinical stages (0-1B) as comparison. Odds for having an anxiety disorder at follow-up followed a linear trend, with higher stages being at higher risk:

at six-year follow-up, presence of anxiety disorders for stage 4B: OR = 11.8 (95%CI: 8.39-16.6), stage 4A: OR = 6.27 (95%CI: 4.10-9.58), stage 3B: OR = 5.60 (95%CI: 3.84-8.15), stage 3A: OR = 5.52 (95%CI: 3.31-9.21), stage 2B: OR = 3.30 (95%CI: 2.25-4.85), and stage 2A: OR = 2.45 (95%CI: 1.48-4.04) (see fig. 2A). This pattern was also present after two-year and four-year follow-up (fig. 2A).

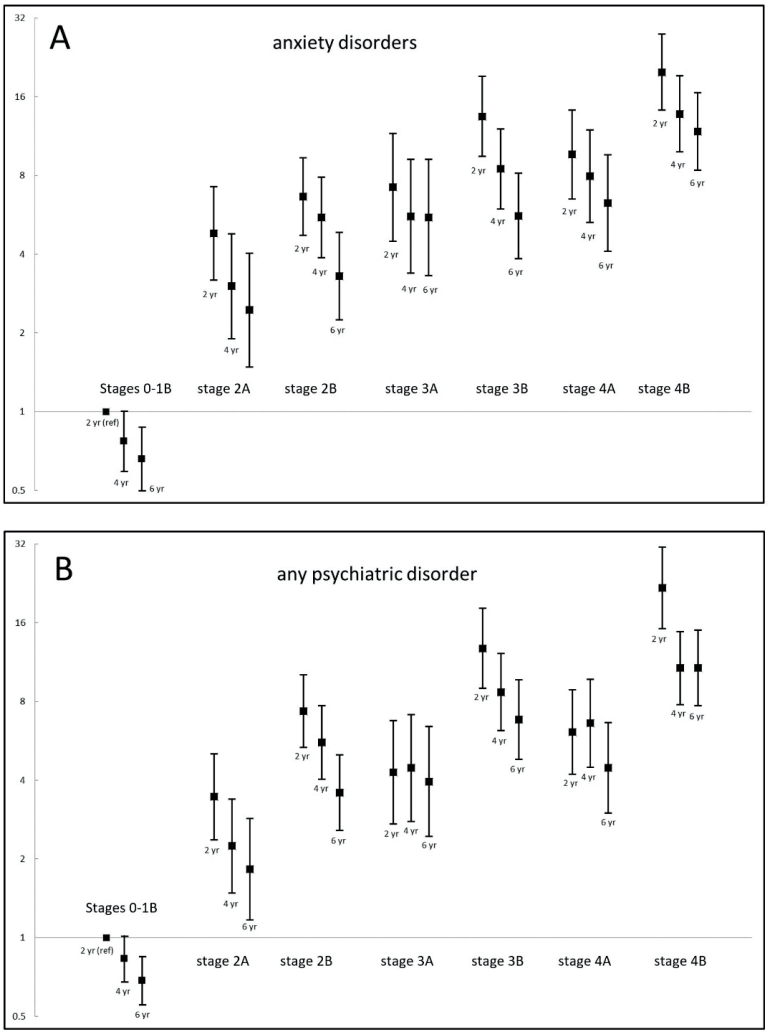


Figure 2. Odds Ratios (bars representing 95% CI) for presence of any anxiety disorder (A) and any psychiatric disorder (B) at 2, 4 and 6-year follow-up with combined subclinical stages 0-1b at 2-year follow-up as reference derived from age-adjusted Generalized Estimating Equations.

Proportions of subjects having any psychiatric disorder at follow-up ranged from 6.5%-8.2% (stage 0) to 68.6%-81.8% (stage 4B). The same pattern as seen with anxiety disorder diagnoses emerged, with the difference that the comorbid stages (2B, 3B, 4B) consistently had the highest odds for having any psychiatric disorder at follow-up (see fig. 2B). For instance, at six-year follow-up, the OR for presence of any psychiatric disorder for stage 4B was 10.7 (7.70-15.0), for stage 4A OR = 4.47 (95%CI: 3.00-6.65), stage 3B OR = 6.82 (95%CI: 4.81-9.67), stage 3A OR = 3.93 (95%CI: 2.44-6.41), stage 2B OR = 3.59 (95%CI: 2.58-5.00), stage 2A OR = 1.83 (95%CI: 1.17-2.85). See eTable 1 for full GEE models.

Subjects in stage 0 had consistently low mean anxiety, depression and disability scores over time. For example, the mean estimated BAI score for stage 0 at baseline = 2.82 (95%CI: 2.16-3.48), at T2 = 2.98 (2.29-3.67), at T4 = 3.04 (2.35-3.74), at T6 = 3.38 (2.67-4.08) (see fig. 3A). The mean anxiety, depression, and disability scores in all other stages were significantly higher in comparison with stage 0 (all $p<0.01$). Estimated means for stages 1B and 2A were not statistically different. Stages 2B, 3A and 4A did not differ significantly with regard to estimated mean anxiety, pathological worrying and social avoidance over time, but these stages all had lower scores over time than stage 3B (see figs. 3A-C). Estimated mean agoraphobic avoidance scores over time were most closely related to successive stages (see fig. 3D). The estimated mean levels of disability and depression over time were significantly higher in the comorbidity stages (2B, 3B, 4B) than in the anxiety-only stages (2A, 3A, 4A) (see fig. 3E-F).

Sensitivity analysis

All included predictors uniquely contributed to the prediction of presence of two-year follow-up anxiety disorders (all p-values <0.001: see e-table 2, model 1). Anxiety severity was the strongest predictor (OR for moderate to severe anxiety = 6.36 (95% CI: 4.00-10.0), while comorbidity had a relatively modest effect: OR = 1.49 (95% CI: 1.19-1.92). However, it should be noted that the effect of comorbidity on longitudinal course existed independently of other predictors. Therefore, the predictors used in the current staging algorithm are all deemed important individual predictors for anxiety course trajectories.

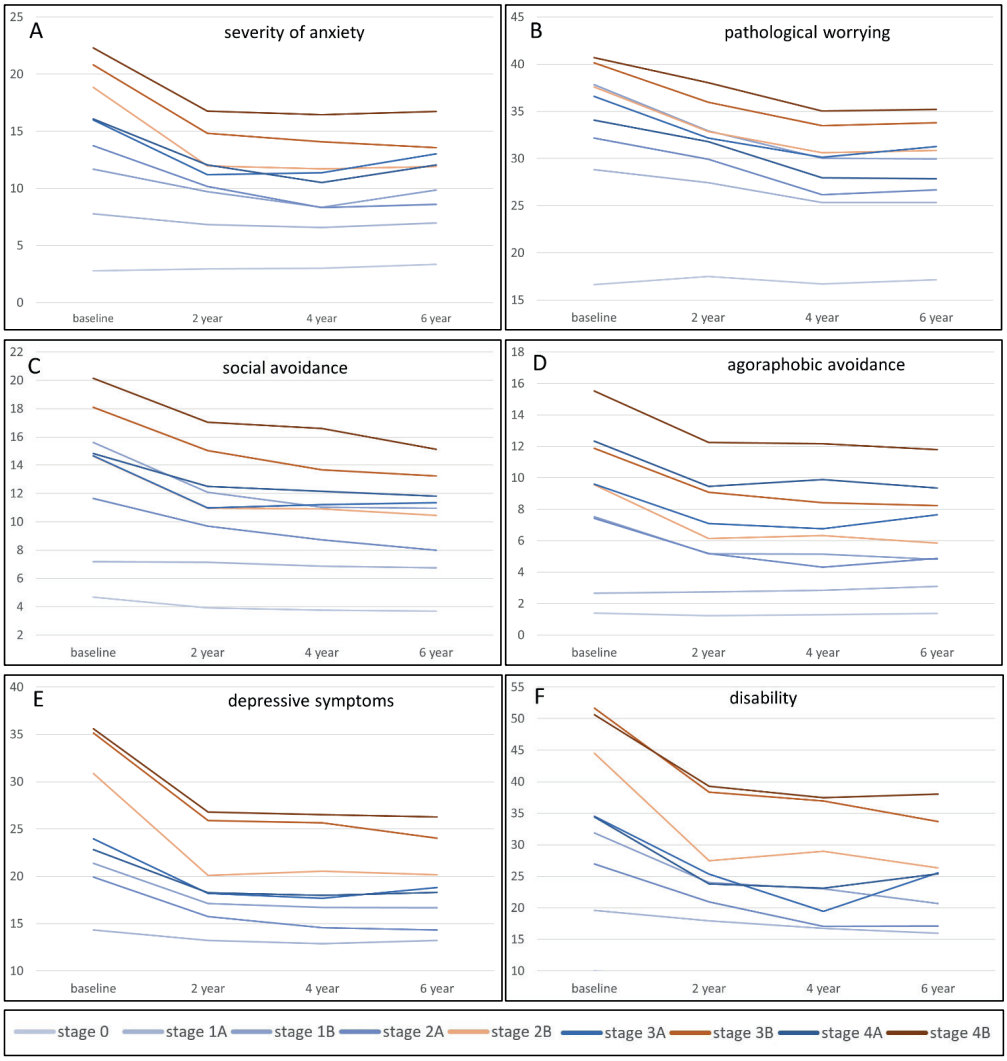


Figure 3. Estimated mean severity of anxiety (panel A), pathological worrying (panel B), social avoidance (panel C), agoraphobic avoidance (panel D), disability (panel F), depression (panel G) per stage at baseline, 2, 4 and 6-year follow-up.

Discussion

In this study, we tailored McGorry's clinical staging model to anxiety disorders and tested its construct and predictive validity in a heterogeneous anxiety disorders sample. First, the presence of construct validity was confirmed by showing that probabilities of unfavorable disease markers increased with each successive stage in the model. This suggests that this staging model is able to distinguish subgroups with increasing levels of disease progression. Second, predictive validity was demonstrated by worsening follow-up outcomes up to six years in higher baseline stages. All associations followed linear trends: severity of anxiety, depression, and disability increased in higher stages. This implies that the process of staging can have value in long-term course prediction in anxiety disorders. It could thus be used as a tool to inform patients about their probable long-term prognosis. Some instructions for use in clinical practice are provided in the supplement. This model could therefore make an important contribution towards the goal of personalized medicine in anxiety disorders.

Different patterns emerged when comparing the different sets of validators: mean values for the anxiety measures followed the successive ordering of stages in the staging model more closely than the mean values of the general psychopathology measures. Subjects in the comorbidity stages (2B, 3B, 4B) had a worse overall longitudinal course compared with those in stages 2A, 3A and 4A. This corroborates the conclusion of other longitudinal studies that presence of comorbidity is associated with poorer long-term outcome.[20] The data suggest that, in all stages, comorbidity impacted the outcome in similar ways: each "B" stage showed worse severity and longitudinal course in comparison with its "A" counterpart. Additionally, in some validators, presence of psychiatric comorbidity seemed to have a higher impact in comparison with anxiety duration. For instance, disability and depressive symptoms over a six year span were most prominent in stages 2B, 3B and 4B, whereas anxiety severity, pathological worrying and social avoidance over a six year span were most impaired in stage 4B followed by 3B.

Limitations

The current study had several limitations. First, associations between the staging model and the validators were not perfectly linear; for instance, stages 3A and 4A showed similar symptom severity at baseline and at follow-up. This could imply that the criteria for stage assignment were not optimal and should be fine-tuned. For instance, in a previous study, threshold social

anxiety disorder in young adults was sometimes assigned to stage 1B in instead of stages 2 and onwards, as social anxiety disorder is hypothesized to be an early stage that will develop into more severe syndromes in a later stage.[10] Second, the current study was limited by the inclusion of only five DSM-IV anxiety disorders. The presence of specific phobias was not assessed, even though it was shown that these disorders may serve as predictors of worse overall longitudinal course.[35] However, the simultaneous assessment of the anxiety disorders that were included is warranted, as these disorders share genetic vulnerability,[36] are highly comorbid,[20] generally show a comparable course,[1] and show diagnostic instability over time.[24,25] Third, from a methodological perspective, the current results might represent an overestimation (i.e. optimism). Applying another external validation dataset or applying a bootstrapping approach might have resulted in more modest estimations on predictive power. However, the lowest EPVs in any of our binary predictive analyses were well above the suggested threshold of 20.[37] As EPVs in our study were high, optimism is likely to be small. Furthermore, it remains unclear whether the current clinical staging model will be associated with underlying pathophysiological processes involved in etiology of anxiety disorders. Possibly, stage assignment criteria need to be refined to reflect underlying disease processes in the future. Finally, one of the major goals of clinical staging models is to derive more targeted interventions that prevent progression across stages of anxiety disorders.[38] However, we were not able to test the applicability of the current model in treatment decision-making, because NESDA is a naturalistic cohort study. This should be a priority in future research as increasing knowledge of effective stage-specific treatments can contribute significantly to the development of personalized medicine.

Future research

In staging models for depressive and bipolar disorders, anxiety disorders are viewed as a nonspecific prodromal phase, which could function as a gateway to development of these end-stage syndromes.[39] Conceptually, staging models for different end-stage syndromes can include similar nonspecific prodromal stages as this is in line with the transdiagnostic assumptions underlying staging models.[4,40] This implies that a person with subthreshold anxiety symptoms is considered to be at stage 1 of both an anxiety disorder and a bipolar disorder staging model. Further research should identify profilers that critically determine the pathways to various end-stage syndromes.

In this tailored staging model, patients with remitted anxiety disorder were assigned to a subclinical stage, making the model bidirectional. Due to the "waxing and waning" longitudinal course of anxiety disorders, such a bidirectional model is most likely to fit best.[41] On the other hand, it is also plausible that patients with multiple episodes of anxiety disorders have a less favorable prognosis, suggesting a one-directional model. Future studies could compare bidirectional to one-directional staging models in anxiety disorders.

In our adaptation of McGorry's clinical staging model, comorbidity was added as "B" substages. The current results suggest that this approach is valid. However, future studies could compare anxiety disorder staging models with and without comorbidity substages, or with a different role in stage assignment for presence of comorbidity, to evaluate this approach further. The current model could be refined by studying anxiety disorder relapses in subjects after remission. Supplementary validation of the current staging model could also be carried out by prospectively applying it to an anxiety disorders cohort, as is done in youth mental health care.[8,42]

Improving longitudinal course predication in anxiety disorders might also be possible using other methodological approaches,[4,43,44] such as machine learning algorithms,[45] or network analysis.[46] The advantage of clinical staging over these alternative approaches, however, is its reliance on simple clinical parameters that both clinicians and patients are familiar with.[7] Additionally, staging models are widely used in other fields of medicine, which improves familiarity of patients and clinicians with these models and will thus aid their implementation of these models.

Conclusion

The present study is the first attempt to tailor a staging model to anxiety disorders. The results show that such a model could be clinically meaningful. In this study, we not only predicted anxiety disorder specific phenomena, but also adapted a transdiagnostic view by predicting "any psychiatric disorder" at two, four and six years. Both approaches were effective. This suggests that, in providing individual course prognosis, not only persistence of anxiety disorders should be considered, but also other disorders such as depressive disorders and substance-use disorders. If these results could be replicated and fine-tuned, clinical stage assignment could improve the diagnostic

process of patients with an anxiety disorder. The results from the current, first study on staging in anxiety disorders are promising. To evolve the field of individualized course prediction and treatment decision-making in anxiety disorders, clinical staging could definitely be the way to go.

REFERENCES

1. Batelaan NM, Rhebergen D, Spinhoven P, van Balkom AJ, Penninx BWJH. Two-Year Course Trajectories of Anxiety Disorders: Do DSM Classifications Matter? *J Clin Psychiatry*. 2014;75(09):985-993. doi:10.4088/JCP.13m08837
2. Bruce SE, Yonkers KA, Otto MW, et al. Influence of Psychiatric Comorbidity on Recovery and Recurrence in Generalized Anxiety Disorder, Social Phobia, and Panic Disorder: A 12-Year Prospective Study. *Am J Psychiatry*. 2005;162(6):1179-1187. doi:10.1176/appi.ajp.162.6.1179
3. van Beljouw IM, Verhaak PF, Cuijpers P, van Marwijk HW, Penninx BW. The course of untreated anxiety and depression, and determinants of poor one-year outcome: a one-year cohort study. *BMC Psychiatry*. 2010;10:86. doi:10.1186/1471-244X-10-86
4. Scott J, Leboyer M, Hickie I, et al. Clinical staging in psychiatry: a cross-cutting model of diagnosis with heuristic and practical value. *Br J Psychiatry*. 2013;202(4):243-245. doi:10.1192/bjp.bp.112.110858
5. Fava GA, Kellner R. Staging: a neglected dimension in psychiatric classification. *Acta Psychiatr Scand*. 1993;87(4):225-230.
6. Cosci F, Fava GA. Staging of mental disorders: systematic review. *Psychother Psychosom*. 2013;82(1):20-34. doi:10.1159/000342243
7. McGorry PD, Hickie IB, Yung AR, Pantelis C, Jackson HJ. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. *Aust N Z J Psychiatry*. 2006;40(8):616-622. doi:10.1111/j.1440-1614.2006.01860.x
8. Hickie IB, Scott EM, Hermens DF, et al. Applying clinical staging to young people who present for mental health care. *Early Interv Psychiatry*. 2013;7(1):31-43. doi:10.1111/j.1751-7893.2012.00366.x
9. Verduijn J, Milaneschi Y, van Hemert AM, et al. Clinical staging of major depressive disorder: an empirical exploration. *J Clin Psychiatry*. 2015;76(9):1200-1208. doi:10.4088/JCP.14m09272
10. Clarke PJF, Hickie IB, Scott E, Guastella AJ. Clinical staging model applied to young people presenting with social anxiety. *Early Interv Psychiatry*. 2012;6(3):256-264. doi:10.1111/j.1751-7893.2012.00364.x
11. Andrews G, Bell C, Boyce P, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for the treatment of panic disorder, social anxiety disorder and generalised anxiety disorder 799453A NP ANZJP ArticlesAndrews et al. RANZCP Guidelines. *Aust New Zeal J Psychiatry*. 2018;52(12):1109-1172. doi:10.1177/0004867418799453

12. Bandelow B, Zohar J, Hollander E, Kasper S, Möller H-J, WFSBP Task force on treatment guidelines for anxiety obsessive-compulsive post-traumatic stress disorders. World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for the pharmacological treatment of anxiety, obsessive-compulsive and post-traumatic stress disorders - First revision. *World J Biol Psychiatry*. 2008;9(4):248-312. doi:10.1080/15622970802465807
13. National Institute for Health and Clinical Excellence. Social Anxiety Disorder: recognition, assessment and treatment. 2013;(May 2013). nice.org.uk/guidance/cg159.
14. Penninx BWJH, Beekman ATF, Smit JH, et al. The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *Int J Methods Psychiatr Res*. 2008;17(3):121-140. doi:10.1002/mpr.256
15. Moreno-Peral P, Conejo-Cerón S, Motrico E, et al. Risk factors for the onset of panic and generalised anxiety disorders in the general adult population: A systematic review of cohort studies. *J Affect Disord*. 2014;168:337-348. doi:10.1016/j.jad.2014.06.021
16. Batelaan NM, Smit F, de Graaf R, van Balkom AJLM, Vollebergh WAM, Beekman ATF. Identifying target groups for the prevention of anxiety disorders in the general population. *Acta Psychiatr Scand*. 2010;122(1):56-65. doi:10.1111/j.1600-0447.2009.01488.x
17. Graaf D, Bijl R V, Smit F, Vollebergh WAM, Spijker J. Risk factors for 12-month comorbidity of mood, anxiety, and substance use disorders: Findings from the Netherlands Mental Health Survey and Incidence Study. *Am J Psychiatry*. 2002;159(4):620-629. doi:10.1176/appi.ajp.159.4.620
18. Macmillan HL, Fleming JE, Streiner DL, et al. Childhood abuse and lifetime psychopathology in a community sample. *Am J Psychiatry*. 2001;158(11):1878-1883. doi:10.1176/appi.ajp.158.11.1878
19. Hyland P, Shevlin M, Elklit A, Christoffersen M, Murphy J. Social, familial and psychological risk factors for mood and anxiety disorders in childhood and early adulthood: a birth cohort study using the Danish Registry System. *Soc Psychiatry Psychiatr Epidemiol*. 2016;51(3):331-338. doi:10.1007/s00127-016-1171-1
20. Lamers F, van Oppen P, Comijs HC, et al. Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). *J Clin Psychiatry*. 2011;72(3):341-348. doi:10.4088/JCP.10m06176blu
21. World Health Organization. *World Health Organization, Composite International Diagnostic Interview (CIDI), Core Version 2.1.* Geneva; 1998.
22. Lyketsos CG, Nestadt G, Cwi J, Heithoff K, et al. The Life Chart Interview: A standardized method to describe the course of psychopathology. *Int J Methods Psychiatr Res*. 1994;4(3):143-155.

23. Warshaw MG, Keller MB, Stout RL. Reliability and validity of the longitudinal interval follow-up evaluation for assessing outcome of anxiety disorders. *J Psychiatr Res*. 1994;28(6):531-545. doi:10.1016/0022-3956(94)90043-4
24. Hovenkamp-Hermelink JHM, Riese H, Van Der Veen DC, Batelaan NM, Penninx BWJH, Schoevers RA. Low stability of diagnostic classifications of anxiety disorders over time: A six-year follow-up of the NESDA study. *J Affect Disord*. 2016;190:310-315. doi:10.1016/j.jad.2015.10.035
25. Scholten WD, Batelaan NM, Penninx BWJH, et al. Diagnostic instability of recurrence and the impact on recurrence rates in depressive and anxiety disorders. *J Affect Disord*. 2016;195:185-190. doi:10.1016/j.jad.2016.02.025
26. Bystritsky A. Treatment-resistant anxiety disorders. *Mol Psychiatry*. 2006;11(9):805-814. doi:10.1038/sj.mp.4001852
27. Iancu SC, Batelaan NM, Zweekhorst MBM, et al. Trajectories of functioning after remission from anxiety disorders: 2-year course and outcome predictors. *Psychol Med*. 2014;44(3):593-605. doi:10.1017/S0033291713001050
28. Beck AT, Steer RA, Carbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clin Psychol Rev*. 1988;8(1):77-100. doi:10.1016/0272-7358(88)90050-5
29. Marks IM, Mathews AM. Brief standard self-rating for phobic patients. *Behav Res Ther*. 1979;17(3):263-267. doi:10.1016/0005-7967(79)90041-X
30. Drost J, Der Does AJW, Antypa N, Zitman FG, Van Dyck R, Spinhoven P. General, specific and unique cognitive factors involved in anxiety and depressive disorders. *Cognit Ther Res*. 2012;36(6):621-633. doi:10.1007/s10608-011-9401-z
31. Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol Med*. 1996;26(3):477-486. doi:10.1017/S0033291700035558
32. Chwastiak LA, Von Korff M. Disability in depression and back pain: Evaluation of the World Health Organization Disability Assessment Schedule (WHO DAS II) in a primary care setting. *J Clin Epidemiol*. 2003;56(6):507-514. doi:10.1016/S0895-4356(03)00051-9
33. Cuzick J. A Wilcoxon-type test for trend. *Stat Med*. 1985;4(1):87-90.
34. Twisk JWR, de Vente W, Apeldoorn AT, de Boer MR. Should we use logistic mixed model analysis for the effect estimation in a longitudinal RCT with a dichotomous outcome variable? *Epidemiol Biostat Public Heal*. 2017;14(3):e12613-1-e12613-8. doi:10.2427/12613
35. Wardenaar KJ, Lim CCW, Al-Hamzawi AO, et al. The cross-national epidemiology of specific phobia in the World Mental Health Surveys. *Psychol Med*. 2017;47:1744-1760. doi:10.1017/S0033291717000174

36. Otowa T, Hek K, Lee M, et al. Meta-analysis of genome-wide association studies of anxiety disorders. *Mol Psychiatry*. 2016;21(10):1391-1399. doi:10.1038/mp.2015.197
37. Austin PC, Steyerberg EW. Events per variable (EPV) and the relative performance of different strategies for estimating the out-of-sample validity of logistic regression models. *Stat Methods Med Res*. 2017;26(2):796-808. doi:10.1177/0962280214558972
38. McGorry PD, Ratheesh A, O'Donoghue B. Early Intervention—an Implementation Challenge for 21st Century Mental Health Care. *JAMA Psychiatry*. 2018;75(6):545-546. doi:10.1001/jamapsychiatry.2018.0621
39. Duffy A. Toward a comprehensive clinical staging model for bipolar disorder: integrating the evidence. *Can J Psychiatry*. 2014;59(12):659-666.
40. Buckholtz JW, Meyer-Lindenberg A. Psychopathology and the human connectome: toward a transdiagnostic model of risk for mental illness. *Neuron*. 2012;74(6):990-1004. doi:10.1016/j.neuron.2012.06.002
41. McGorry PD, Purcell R, Hickie IB, Yung AR, Pantelis C, Jackson HJ. Clinical staging: a heuristic model for psychiatry and youth mental health. *Med J Aust*. 2007;187(7 Suppl):S40-2.
42. Purcell R, Jorm AF, Hickie IB, et al. Demographic and clinical characteristics of young people seeking help at youth mental health services: baseline findings of the Transitions Study. *Early Interv Psychiatry*. March 2014. doi:10.1111/eip.12133
43. McGorry PD. Early clinical phenotypes, clinical staging, and strategic biomarker research: building blocks for personalized psychiatry. *Biol Psychiatry*. 2013;74(6):394-395. doi:10.1016/j.biopsych.2013.07.004
44. Hasler G, Drevets WC, Manji HK, Charney DS. Discovering Endophenotypes for Major Depression. *Neuropsychopharmacology*. 2004;29(10):1765-1781. doi:10.1038/sj.npp.1300506
45. Iniesta R, Stahl D, McGuffin P. Machine learning, statistical learning and the future of biological research in psychiatry. *Psychol Med*. 2016;46(May):2455-2465. doi:10.1017/S0033291716001367
46. Borsboom D, Cramer OJ, Cramer AOJ. Network analysis: an integrative approach to the structure of psychopathology. *Annu Rev Clin Psychol*. 2013;9:91-121. doi:10.1146/annurev-clinpsy-050212-185608

Supplementary materials

Additional methods

Sample

The main aim of the Netherlands Study of Depression and Anxiety (NESDA) is unravelling the factors that determine the natural course of depression and anxiety. NESDA used a stratified sampling approach, including subjects from the community, primary health care and specialized mental health care in different regions in The Netherlands. Exclusion criteria included a primary diagnosis of other psychiatric disorders, such as psychotic disorders, bipolar disorder, posttraumatic stress disorder, or obsessive-compulsive disorder. Those with insufficient control of the Dutch language were also excluded. At baseline, 2,981 subjects were included. Measurements included sociodemographics, clinical characteristics, psychological assessments, biological assessments and structured psychiatric interviews assessing DSM-IV diagnoses. The Ethical Committee of participating sites approved of the study design and all subjects provided written informed consent. An extensive description of the study design is provided elsewhere.[1]

We used data from baseline measurements (2004-2007) and from 2, 4 and 6-year follow-up measurements. As clinical staging tools were designed to assign at-risk subjects and subjects seeking help for mental health issues to clinical stages, but not healthy controls [2,3], we included symptomatic subjects (presence of any DSM-IV anxiety disorder) and asymptomatic subjects at risk for developing an anxiety disorder from the NESDA baseline sample. We operationalized this as either having a lifetime diagnosis of an anxiety disorder (n=1,772), having been exposed to childhood trauma (n=1,442), or having a positive family history for psychiatric disorders (n=2,585), as these factors are important predictors of future anxiety disorder development.[4-7] In this manner, 156 healthy controls who did not have risk factors were excluded.

Non-response was associated with lower years of education ($p < 0.001$, Cohen's $d=0.29$), childhood trauma ($p<0.001$) and lifetime history of anxiety disorders ($p<0.001$). In comparison with clinical stage 0, non-response was significantly greater in stages 2A, 2B, 3B, 4A, and 4B, with ORs varying between 2.37 (stage 2B) and 2.75 (stage 3B). This introduced a small bias towards the null hypothesis. Non-response was not associated with gender, age, and family history for psychiatric disorders.

Measurements

Staging model

Our main independent variable was a new staging model aimed at staging subjects with anxiety disorders. McGorry's generic model [2,3] was adapted to derive this staging model. In their generic model, stages are described within a transdiagnostic conceptual framework. However, in order to assess the staging model empirically, we had to deviate from their model in two important ways. First, we chose to investigate anxiety disorders as primary end-stage syndromes. Therefore, in our sample we excluded other psychiatric end-stage syndromes if anxiety was absent. Second, in order to assign subjects to stages, we needed to operationalize the proposed criteria in a quantitative manner (see below).

The different stages in this model were stage 0 (asymptomatic), 1A (nonspecific symptoms), 1B (attenuated syndromes), 2A (discrete disorder), 2B (discrete disorder with comorbidity), 3A (intermittent symptoms), 3B (intermittent symptoms with comorbidity), 4A (chronic symptoms), 4B (chronic symptoms with comorbidity). Various assignment criteria were used to assign subjects to our adaptation of McGorry's model (see Table 1 for an overview).

Subclinical stages

Subthreshold symptoms were permitted in these stages. This was done as McGorry's staging model is aimed at at-risk populations, not healthy individuals. Lifetime anxiety disorder diagnoses were assessed using the WHO-Composite International Diagnostic Interview (CIDI, version 2.1). The CIDI is a structured interview that classifies according to DSM-IV criteria.[8,9] Childhood trauma was assessed in a face-to-face interview, using the Childhood Trauma Interview (CTI), which was used before in the Netherlands Mental Health Survey and Incidence Study.[10] For the purpose of this study, childhood trauma was deemed present when subjects experienced at least one instance of emotional neglect, psychological abuse, physical abuse or sexual abuse before the age of sixteen. Family history for anxiety disorders, MDD and alcohol dependency was assessed in a structured manner by asking subjects about the possible symptomatology in family members using the family tree method [11].

Distributions across the subclinical stages were made by using severity of anxiety, worrying and avoidance behaviors, based on symptom severity cut-off scores: stage 0 for low symptom severity, stage 1A for mild to moderate symptom severity, and stage 1B for moderate to severe symptom severity.

Severity of anxiety was measured with the Beck Anxiety Inventory (BAI), a 21-item self-report questionnaire.[12]. Although the BAI is not a diagnostic tool, there is evidence that it may discriminate between subjects with and without anxiety disorders. A score below a cut-off value of 10 resulted in a negative predictive power of 0.97 for not having an anxiety disorder, whereas a score above the cut-off value of 30 resulted in a positive predictive value of 0.40 for having an anxiety disorder.[13]. We regarded a score of <10 as low anxiety severity, and one of >30 as moderate to severe anxiety.

Severity of avoidance was measured with the Fear Questionnaire (FQ), a 15-item self-report questionnaire.[14] The FQ reflects the rate of avoidance in three subscales: agoraphobia (FQ Ag), social phobia (FQ So) and blood injury phobia. The third subscale was omitted for the purpose of this study, as specific phobias are not within the scope of this study. Different cut-off values were used in the FQ subscales. A study by Schadé et al used more lenient cut-off values (<15 for FQ Ag and <12 for FQ So) in comparison with the study of van Zuuren (<19 for FQ Ag and <18 for FQ So).[15,16] For the purpose of this study, we combined the cut-off scores: low avoidance severity was defined by scores below the cut-off values by Schadé et al. and moderate to severe avoidance was defined by scores above those by van Zuuren.

Severity of pathological worrying was assessed using the 11-item self-report version of the Penn State Worry Questionnaire (PSWQ).[17,18] The PSWQ was repeatedly shown to have good reliability and validity.[17,19] A total score below 24 indicates low levels of pathological worrying, whereas a score above 39 indicates severe levels of pathological worrying.[20] In summary, the three subclinical stages were defined as follows:

low:	BAI<10	and PSWQ<24 and FQ (Ag)<15	and FQ (So) <12
mild to moderate:	BAI 10<30	or PSWQ 24<39 or FQ (Ag) 15<19	or FQ (So) 12<18
moderate to severe:	BAI ≥30	or PSWQ ≥39 or FQ (Ag) ≥19	or FQ (So) ≥18.

Clinical stages

To account for differences in number of months in the current calendar year that were assessed, number of months were recalculated into proportional scores. These proportional scores reflect duration of anxiety disorders: 'not chronic' (<30% of months), 'intermittent' (30-80% of months), and 'chronic' (>80% of months). These values were chosen to reflect different levels of chronicity for anxiety disorders, in order to conform to Hickie's descriptions for each different stage.

If current MDD, dysthymia or alcohol dependency was present, subjects were assigned to the 'b' stages 2B, 3B and 4B, otherwise, they were assigned to the 'a' stages 2A, 3A and 4A. The decision to include the presence of comorbid psychiatric disorders as a sub-stage criterium, or 'profiler', was informed by the high levels of comorbidity between anxiety disorders with other common mental disorders, such as depressive disorders and alcohol dependence. As many as 63% of persons with a current anxiety disorder also have a current depressive disorder.[21] Those with comorbid disorders are also burdened by higher degrees of anxiety severity, disability and other markers of poor outcome,[21,22] making anxiety disorders and depressive disorders highly intertwined. As comorbidity is such an essential aspect of anxiety disorders, and as staging models aim for a transdiagnostic approach, it is in our view essential to include comorbidity into the model. Even though we only assessed the staging model at baseline, in our staging model it is theoretically possible to decrease in clinical stage after follow-up (i.e.: bidirectional). See Table 1 for all assignment criteria and figure 1 for a flowchart of inclusion and stage assignment.

Construct validators

In clinical staging, it is assumed that patients with more advanced disease are burdened with greater levels of symptomatology. As symptoms in anxiety disorders often overlap with those in other common mental disorders, we distinguished two sets of validators: anxiety validators and general psychopathology validators. The set of anxiety validators consisted of four baseline variables: severity of anxiety (BAI), severity of agoraphobic avoidance behaviors (FQ), severity of social phobia (FQ), and severity of pathological worrying (PSWQ). Incorporating general psychopathology validators is in line with earlier research establishing the need to take functional recovery into account as an important outcome measure, not only remission status of anxiety disorders.[23] It also takes diagnostic instability within anxiety disorders and between anxiety disorders and depressive

disorders into account.[24] General psychopathology validators were severity of depressive symptoms and levels of disability. Depressive symptoms were measured with the Inventory of Depressive Symptomatology-SR (IDS), a self-report questionnaire on severity of depression;[25] levels of disability were measured with the WHO-Disability Assessment Schedule (WHO-DAS II), a 36-item self-report questionnaire measuring levels of disability.[26]

Predictive validators

Follow-up measurements for the construct validators were used as predictive validators. Additionally, presence of DSM-IV anxiety disorders (CIDI, see above) and presence of any psychiatric disorder (either anxiety disorder, MDD, dysthymia or alcohol dependency; CIDI, see above) were used as predictive validators. These validators were measured at 2, 4 and 6-year follow-up. It was assumed that worse follow-up outcomes on DSM classifications reflect a worse longitudinal course, as these represent persistent or relapsed disease. Depressive disorders and alcohol dependency were included as general psychopathology validators because longitudinal anxiety course is known to show high levels of comorbidity and diagnostic instability: during progression of or even after remission of an anxiety disorder other psychiatric disorders frequently occur,[21,24,27] which indicates a poor outcome. In this way it was possible to evaluate the model in a transdiagnostic manner. As the main goal of this model is to predict longitudinal course of anxiety, *anxiety* course was regarded our primary longitudinal outcome and *general psychopathology* course a secondary longitudinal outcome.

Sample characteristics

A number of demographic and clinical characteristics were assessed in order to describe our sample. Age, gender and education level were recorded at baseline by using a demographics questionnaire. Whether subjects received psychotherapy was assessed with the Perceived Need for Care Questionnaire (PNCQ). The PNCQ is a semi-structured interview assessing the current mental health care use and patient's perceived need for mental health care interventions,[28] and was slightly altered for use in the NESDA study.[29] We derived data on proportion of subjects currently receiving psychotherapy, as defined by at least 5 consultations with a psychologist, psychiatrist, psychotherapist or other mental health care worker during the previous 6 months, as done before.[6] The use of psychotropic medication (tricyclic antidepressants (TCAs), selective serotonin reuptake inhibitors (SSRIs), selective noradrenaline reuptake inhibitors (SNRIs), monoamine oxidase

inhibitors (MAO-i), and other antidepressants) was assessed via inspection of medication containers and via classification according to the Anatomical Therapeutic Chemical (ATC) classification. For the purpose of this study, we extracted data on frequent use (>50% of days) of any antidepressant. We reported proportions of having any form of treatment (antidepressants or psychotherapy).

Statistical analyses

Predictive validity analyses

The GEE models were specified by a binomial distribution for the dependent variable, a logit link function, an exchangeable correlation structure, and were fitted to the longitudinal data of wave T2, T4 and T6 to estimate effects of stage (indicators of stages 2A, 2B, 3A, 3B, 4A and 4B with the reference group that consisted of the combined stages 0, 1A and 1B), time-points (discrete indicators of T4 and T6 with the reference time point T2) and all stage*time-point interactions, adjusting for baseline age. This approach nullifies selective attrition effects at follow-up by estimating effects for each stage, each timepoint, and stage*time interactions in one model. Full model details are provided in this supplement (Supplementary Table 7). For continuous validators linear mixed models were performed. A 2-level model with a random intercept was employed in which subject-identifier was the highest level, this way correlations among subjects across follow-up measurements were taken into account. For continuous outcomes the use of an LMM was preferred over the use of GEE because of its superiority in handling incomplete datasets.[30] We calculated mean values with 95% confidence intervals for each stage at each time-point (data for confidence intervals available at request).

Sensitivity analyses

In order to test whether all variables that were included in the staging model indeed had a *unique* contribution to the longitudinal outcome prediction, an analysis in which all variables used in stage assignment were included to a logistic regression model predicting two-year presence of anxiety disorders. In this model we included categorical predictors severity of anxiety, according to the cut-off values we used in the staging model. So, like in the staging model, we categorized into three groups: low severity, mild to moderate severity and moderate to severe. The second categorical predictor in this logistic regression model was duration of anxiety and/or avoidance symptoms. For this categorical predictor with mutually exclusive groups, the same cut-off values that were used in the staging model were used:

No anxiety symptoms: subjects without anxiety disorders

No chronic duration: <30% of the months presence of anxiety and avoidance symptoms during the last four years before baseline

Intermittent duration: 30-80% of the months presence of anxiety or avoidance symptoms (highest value counts) during the last four years before baseline

Chronic duration: >80% of the months presence of anxiety or avoidance symptoms (highest value counts) during the last four years before baseline

The other predictors in the logistic model were presence of comorbidity (six month diagnoses of major depressive disorder, dysthymia, or alcohol dependency), age, gender and education level (years). The combination of these variables represent the totality of information used for stage assignment. Logistic regression analysis on presence of two-year anxiety disorders were performed in order to assess whether all variables used in stage assignment have a contribution to course prediction. This was done in a multivariable model, to assess each variables' unique contribution. We hypothesized that all variables that were used in stage assignment were significant predictors longitudinal course, justifying our decision of using them in stage assignment.

Secondly, logistic regression analysis which included the clinical stages as categorical predictor variables were performed. In comparison to the first model all other predictors were removed, except for age, gender and education level. We hypothesized that if the second model properties would approximate the properties in Model 1, applying the stratified staging model is valid.

Supplementary results

In eTable 1, the full GEE model can be found.

As can be seen in eTable 2 (Model 1), all included predictors had a unique contribution to prediction of presence of two-year follow-up anxiety disorder (all p-values <0.001), while sociodemographic variables did not predict anxiety course. Anxiety severity was the strongest predictor (OR for moderate to severe anxiety= 6.36 (95% CI: 4.00-10.0), while comorbidity had a relatively modest effect: OR=1.49 (95% CI: 1.19-1.92). However, it should be noted that the effect of comorbidity on longitudinal course exists independently from anxiety severity or duration. Therefore, the presence of comorbidity should be considered an important individual predictor for anxiety course.

In the second model, all clinical stages predicted two year presence of anxiety disorders (all p-values <0.001) and ORs ranging from 6.18 (stage 1A) to 74.6 (stage 4B). Sociodemographic variables did not predict longitudinal course. The c-statistic, which is statistically similar to the Area under the curve for predictions based on this logistic regression model, is very comparable between these two models: 0.821 (95% CI: 0.802-0.839) for Model 1 and 0.815 (95% CI: 0.797-0.834) for model 2. Overall, these c-statistic values indicate a good fit (>0.80). The lack of difference between the two models indicates that a stratified approach does not lead to loss of predictive power.

eTable 1. Odds Ratio's estimated from GEE models for presence of any anxiety disorder and for presence of any psychiatric disorder.

	Any anxiety disorder			Any psychiatric disorder		
	OR	**95%-CI**	**p-value**	**OR**	**95%-CI**	**p-value**
Stage						
stage 0,1A,1B	1.00			1.00		
stage 2A	4.70	(3.11, 7.09)	<0.001	3.46	(2.37, 5.06)	<0.001
stage 2B	6.53	(4.64, 9.19)	<0.001	7.34	(5.34, 10.11)	<0.001
stage 3A	7.13	(4.43, 11.46)	<0.001	4.29	(2.72, 6.76)	<0.001
stage 3B	13.38	(9.43, 18.99)	<0.001	12.78	(8.99, 18.17)	<0.001
stage 4A	9.69	(6.54, 14.33)	<0.001	6.11	(4.21, 8.87)	<0.001
stage 4B	19.97	(14.28, 27.92)	<0.001	21.76	(15.20, 31.14)	<0.001
Time						
2 years	1.00			1.00		
4 years	0.77	(0.59, 1.01)	0.055	0.83	(0.68, 1.01)	0.069
6 years	0.66	(0.50, 0.87)	0.004	0.68	(0.55, 0.85)	<0.001
Stage by Time interaction terms						
stage2A*2years	0.81	(0.48, 1.39)	0.449	0.78	(0.48, 1.26)	0.308
stage2A*4years	0.77	(0.44, 1.37)	0.373	0.77	(0.47, 1.28)	0.317
stage2B*2years	1.08	(0.70, 1.65)	0.737	0.92	(0.62, 1.35)	0.659
stage2B*4years	0.75	(0.48, 1.19)	0.224	0.71	(0.48, 1.07)	0.100
stage3A*2years	1.00	(0.56, 1.81)	0.991	1.25	(0.72, 2.18)	0.422
stage3A*4years	1.16	(0.63, 2.13)	0.634	1.35	(0.76, 2.38)	0.305
stage3B*2years	0.82	(0.53, 1.26)	0.354	0.82	(0.54, 1.24)	0.346
stage3B*4years	0.63	(0.40, 0.99)	0.047	0.78	(0.51, 1.20)	0.259
stage4A*2years	1.07	(0.66, 1.74)	0.784	1.30	(0.82, 2.05)	0.259
stage4A*4years	0.99	(0.59, 1.64)	0.958	1.07	(0.67, 1.71)	0.787
stage4B*2years	0.90	(0.59, 1.36)	0.607	0.59	(0.39, 0.90)	0.013
stage4B*4years	0.90	(0.59, 1.38)	0.625	0.72	(0.47, 1.10)	0.133
Age/10	0.95	(0.89, 1.02)	0.138	1.00	(0.95, 1.07)	0.895
Constant	0.13	(0.10, 0.19)	<0.001	0.21	(0.15, 0.28)	<0.001
QIC – null model		7916.282			9345.107	
QIC – model with ICS		5305.043			9051.020	
QIC – model with ECS		5304.766			6050.778	

QIC = model fit in terms of Quasi-likelihood under the Independence model Criterion, ICS = Independent Correlation Structure, ECS = Exchangeable Correlation Structure (i.e. the model for which the parameters are presented)

eTable 2. Odds Ratio's estimated from logistic regression analysis models for presence of any anxiety disorder at two year follow-up, comparing individual assignment criteria with the staging model.

	Model 1				Model 2			
	OR	**95% CI**	**df**	**p**	**OR**	**95% CI**	**df**	**p**
constant	0.04	na	1	<0.001	0.04	na	1	<0.001
Age	1.00	(0.99, 1.01)	1	0.72	1.00	(0.99, 1.01)	1	0.65
Gender (ref = male)	1.19	(0.94, 1.52)	1	0.15	1.25	(0.98, 1.59)	1	0.07
Education level	0.98	(0.95, 1.01)	1	0.20	0.98	(0.94, 1.01)	1	0.15
Severity of anxiety:			2	<0.001				
Low (ref)								
Mild to moderate	3.66	(2.32, 5.79)	1	<0.001				
Moderate to severe	6.36	(4.00, 10.1)	1	<0.001				
Anxiety/ avoidance duration:			3	<0.001				
No anxiety disorder (ref)								
Duration <30%	2.57	(1.81, 3.64)	1	<0.001				
Duration 30-80%	4.39	(3.04, 6.34)	1	<0.001				
Duration >80%	5.81	(4.10, 8.23)	1	<0.001				
Presence of comorbidity	1.49	(1.16, 1.92)	1	<0.001				
Clinical staging model								
Stage 0 (ref)							8	<0.001
Stage 1A					6.18	(3.36, 11.4)		<0.001
Stage 1B					9.36	(4.84, 18.1)		<0.001
Stage 2A					18.4	(9.67, 34.9)		<0.001
Stage 2B					24.5	(13.5, 44.6)		<0.001
Stage 3A					27.3	(13.8, 54.0)		<0.001
Stage 3B					51.8	(28.3, 94.7)		<0.001
Stage 4A					35.8	(19.1, 67.2)		<0.001
Stage 4B					74.6	(41.1, 135.4)		<0.001
Model properties:								
c-statistic	0.821	(0.802, 0.839)		<0.001	0.815	(0.797, 0.834)		<0.001

[1]Severity of anxiety was defined by applying validated cut-off values for Beck's anxiety Inventory (BAI), Penn State Worry Questionnaire (PSWQ), Fear Questionnaire agoraphobic avoidance (FQ (Ag)), and Fear Questionnaire agoraphobic avoidance (FQ (Ag)).

[2] low:	BAI<10	and PSWQ<24	and FQ (Ag)<15	and FQ (So)<12
[3] mild to moderate:	BAI 10<30	or PSWQ 24<39	or FQ (Ag) 15<19	or FQ (So) 12<18
[4] moderate to severe:	BAI ≥30	or PSWQ ≥39	or FQ (Ag) ≥19	or FQ (So) ≥18.

Instructions for use in clinical practice

The generic, heuristic staging model that was operationalized and validated in the current study is available elsewhere [3]. In order to apply the current model in clinical practice, a regular clinical assessment should be made by the practitioner. In this assessment, three steps should be followed.

1. In persons presenting with anxiety symptoms, first *assess the presence of DSM 5 anxiety disorders* in accordance with guidelines. If any anxiety disorder is present, the patient will be assigned to the clinical stages 2, 3 or 4. If no anxiety disorder is present, the person will be assigned to the subclinical stage (0, 1A, 1B).

2. For patients with a DSM anxiety disorder diagnosis, *assess the duration of anxiety disorder symptoms.*
   - If the anxiety disorder symptoms were present for at least 4 out the previous five years, the patient is assigned to stage 4.

- If the anxiety disorder symptoms were present for at least 1.5 out of the last five years, but not more than four out of the five previous years, the patient is assigned to stage 3.

- If the symptoms of the disorder were present in less than 1.5 years during the last five years, the patient is assigned to stage 2. *NOTE*: for research purposes, in our study the life chart inventory was used.[31] This approach optimizes this part of the assessment by using affectively laden personal memory anchors. It was shown previously that these personal memory anchors aid recollection of previous symptoms.[32] It is therefore recommended to use a memory anchor method, like the Life chart method to assess the previous duration of anxiety disorder symptoms. For patients without DSM diagnosis, *assess the severity of anxiety symptoms*.

- If the presenting symptom is social anxiety, use a dedicated social phobia rating scale to assess severity of symptoms. We used the social phobia subscale derived from the Fear Questionnaire but other measurement scales can be used as well, e.g. Leibowitz Social Anxiety Scale (LSAS), or the Social Interaction Anxiety Scale (SIAS). However, in the current paper we did not examine cut-off values for these measurement scales, so therefore we advise to use the Fear Questionnaire.

- If the presenting symptom is generalized anxiety, use a dedicated rating scale. For instance the Beck Anxiety Inventory (BAI) or the Penn State Worry Questionnaire (PSWQ). Proposed cut-off values are presented below.

- If the presenting symptom is panic attacks, use a dedicated rating scale. We used the BAI and Fear Questionnaire (agoraphobia subscale), for which reference values are presented below. Alternatively, the Panic Disorder Severity Scale (PDSS) can be used, but for these, no cut-off values were examined and therefore we advise to use the Fear Questionnaire.

Proposed cut-off values for measurement instruments used in our study:

Stage 0:	BAI<10	PSWQ<24	FQ (Ag)<15	FQ (So) <12
Stage 1A:	BAI 10<30	PSWQ 24<39	FQ (Ag) 15<19	FQ (So) 12<18
Stage 1B:	BAI ≥30	PSWQ ≥39	FQ (Ag) ≥19	FQ (So) ≥18.

3. In clinical stages, *assess the presence of comorbid psychiatric disorders* according to guidelines. If dysthymia, (uni- or bipolar) depressive disorders, alcohol use disorder, posttraumatic stress disorders, obsessive compulsive disorders, personality disorders or psychotic disorders are present, assign patients to B stages (2B, 3B, 4B); otherwise, assign patients to A stages (2A, 3A, 4A).

References

1. Penninx BWJH, Beekman ATF, Smit JH, et al. The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *Int J Methods Psychiatr Res*. 2008;17(3):121-140. doi:10.1002/mpr.256
2. McGorry PD, Hickie IB, Yung AR, Pantelis C, Jackson HJ. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. *Aust N Z J Psychiatry*. 2006;40(8):616-622. doi:10.1111/j.1440-1614.2006.01860.x
3. Hickie IB, Scott EM, Hermens DF, et al. Applying clinical staging to young people who present for mental health care. *Early Interv Psychiatry*. 2013;7(1):31-43. doi:10.1111/j.1751-7893.2012.00366.x
4. Moreno-Peral P, Conejo-Cerón S, Motrico E, et al. Risk factors for the onset of panic and generalised anxiety disorders in the general adult population: A systematic review of cohort studies. *J Affect Disord*. 2014;168:337-348. doi:10.1016/j.jad.2014.06.021
5. Batelaan NM, Rhebergen D, Spinhoven P, van Balkom AJ, Penninx BWJH. Two-Year Course Trajectories of Anxiety Disorders: Do DSM Classifications Matter? *J Clin Psychiatry*. 2014;75(09):985-993. doi:10.4088/JCP.13m08837
6. Scholten WD, Batelaan NM, van Balkom AJ, Penninx BW, Smit JH, Van Oppen P. Recurrence of anxiety disorders and its predictors. *J Affect Disord*. 2013;147(1-3):180-185. doi:10.1016/j.jad.2012.10.031
7. Craske MG, Stein MB, Eley TC, et al. Anxiety disorders. *Nat Rev Dis Prim*. 2017;3:17-24. doi:10.1038/nrdp.2017.24
8. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR Fourth Edition (Text Revision)*.; 2000.
9. World Health Organization. *World Health Organization, Composite International Diagnostic Interview (CIDI), Core Version 2.1.* Geneva; 1998.
10. Graaf D, Bijl R V, Smit F, Vollebergh WAM, Spijker J. Risk factors for 12-month comorbidity of mood, anxiety, and substance use disorders: Findings from the Netherlands Mental Health Survey and Incidence Study. *Am J Psychiatry*. 2002;159(4):620-629. doi:10.1176/appi.ajp.159.4.620
11. Fyer AJ, Weissman MM. Genetic linkage study of panic: Clinical methodology and description of pedigrees. *Am J Med Genet*. 1999;88(2):173-181. doi:10.1002/(SICI)1096-8628(19990416)88:2<173::AID-AJMG15>3.0.CO;2-#
12. Beck AT, Steer RA, Carbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clin Psychol Rev*. 1988;8(1):77-100. doi:10.1016/0272-7358(88)90050-5

13. Kabacoff RI, Segal DL, Hersen M, Van Hasselt VB. Psychometric properties and diagnostic utility of the Beck Anxiety Inventory and the State-Trait Anxiety Inventory with older adult psychiatric outpatients. *J Anxiety Disord*. 1997;11(1):33-47. doi:10.1016/S0887-6185(96)00033-3
14. Marks IM, Mathews AM. Brief standard self-rating for phobic patients. *Behav Res Ther*. 1979;17(3):263-267. doi:10.1016/0005-7967(79)90041-X
15. Van Zuuren FJ. The Fear Questionnaire. Some data on validity, reliability and layout. *Br J Psychiatry*. 1988;153(NOV.):659-662. doi:10.1192/bjp.153.5.659
16. Schade A, Marquenie LA, van Balkom AJLM, et al. The effectiveness of anxiety treatment on alcohol-dependent patients with a comorbid phobic disorder: a randomized controlled trial. *Alcohol Clin Exp Res*. 2005;29(5):794-800.
17. Meyer TJ, Miller ML, Metzger RL, Borkovec TD. Development and validation of the penn state worry questionnaire. *Behav Res Ther*. 1990;28(6):487-495. doi:10.1016/0005-7967(90)90135-6
18. Drost J, Der Does AJW, Antypa N, Zitman FG, Van Dyck R, Spinhoven P. General, specific and unique cognitive factors involved in anxiety and depressive disorders. *Cognit Ther Res*. 2012;36(6):621-633. doi:10.1007/s10608-011-9401-z
19. Fresco DM, Heimberg RG, Mennin DS, Turk CL. Confirmatory factor analysis of the Penn State Worry Questionnaire. *Behav Res Ther*. 2002;40(3):313-323. doi:10.1016/S0005-7967(00)00113-3
20. Van Der Heiden C, Muris P, Bos AER, Van Der Molen H, Oostra M. Normative data for the Dutch version of the Penn State Worry Questionnaire. *Neth J Psychol*. 2009;65(2):69-75. doi:10.1007/BF03080129
21. Lamers F, van Oppen P, Comijs HC, et al. Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). *J Clin Psychiatry*. 2011;72(3):341-348. doi:10.4088/JCP.10m06176blu
22. Penninx BWJH, Nolen WA, Lamers F, et al. Two-year course of depressive and anxiety disorders: results from the Netherlands Study of Depression and Anxiety (NESDA). *J Affect Disord*. 2011;133(1-2):76-85. doi:10.1016/j.jad.2011.03.027
23. Iancu SC, Batelaan NM, Zweekhorst MBM, et al. Trajectories of functioning after remission from anxiety disorders: 2-year course and outcome predictors. *Psychol Med*. 2014;44(3):593-605. doi:10.1017/S0033291713001050
24. Scholten WD, Batelaan NM, Penninx BWJH, et al. Diagnostic instability of recurrence and the impact on recurrence rates in depressive and anxiety disorders. *J Affect Disord*. 2016;195:185-190. doi:10.1016/j.jad.2016.02.025
25. Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol Med*. 1996;26(3):477-486. doi:10.1017/S0033291700035558

26. Chwastiak LA, Von Korff M. Disability in depression and back pain: Evaluation of the World Health Organization Disability Assessment Schedule (WHO DAS II) in a primary care setting. *J Clin Epidemiol*. 2003;56(6):507-514. doi:10.1016/S0895-4356(03)00051-9
27. Hovenkamp-Hermelink JHM, Riese H, Van Der Veen DC, Batelaan NM, Penninx BWJH, Schoevers RA. Low stability of diagnostic classifications of anxiety disorders over time: A six-year follow-up of the NESDA study. *J Affect Disord*. 2016;190:310-315. doi:10.1016/j.jad.2015.10.035
28. Meadows G, Harvey C, Fossey E, Burgess P. Assessing perceived need for mental health care in a community survey: development of the Perceived Need for Care Questionnaire (PNCQ). *Soc Psychiatry Psychiatr Epidemiol*. 2000;35(9):427-435. doi:10.1007/s001270050260
29. Prins MA, Verhaak PFM, van der Meer K, Penninx BWJH, Bensing, Jozien M. Primary care patients with anxiety and depression: Need for care from the patient's perspective. *J Affect Disord*. 2009;119(1-3):163-171. doi:10.1016/j.jad.2009.03.019
30. Twisk JWR. *Applied Longitudinal Analysis for Epidemiology.* 2nd ed. New York: Cambridge University Press; 2013.
31. Lyketsos CG, Nestadt G, Cwi J, Heithoff K, et al. The Life Chart Interview: A standardized method to describe the course of psychopathology. *Int J Methods Psychiatr Res*. 1994;4(3):143-155.
32. Keller MB, Lavori PW, Friedman B, et al. The Longitudinal Interval Follow-up Evaluation: A Comprehensive Method for Assessing Outcome in Prospective Longitudinal Studies. *Arch Gen Psychiatry*. 1987. doi:10.1001/archpsyc.1987.01800180050009

Chapter 6.

Evaluating a dimensional approach to treatment resistance in anxiety disorders: a two-year follow-up study

Journal of Affective Disorders Reports (2021)

Wicher A. Bokma
Neeltje M. Batelaan
Brenda W.J.H. Penninx
Anton J.L.M. van Balkom

Abstract

Background

Treatment resistance in anxiety disorders (TR-AD) has been previously defined by failed prior treatments and by various clinical aspects, but the impact of these aspects on course during subsequent treatments was never studied. Moreover, TR-AD was never studied using a dimensional approach. This study validated aspects of TR-AD and examined whether a TR-AD score was related to two-year course during treatment.

Methods

From the NESDA cohort, anxiety disorder patients who subsequently received treatment were selected (n=679). Literature-derived aspects of TR-AD at baseline included anxiety severity, functional impairments, psychiatric comorbidity, duration, and previous treatments. These were combined into a dimensional TR-AD score. Individual aspects of TR-AD and the TR-AD score were linked to anxiety disorder persistence at two-year follow-up using logistic regression analyses. Predictive properties for the TR-AD score were assessed.

Results

Current symptom severity, psychiatric comorbidity, functional impairments and previous duration of symptoms were closely associated with two-year anxiety disorder persistence, while treatment history was not. The TR-AD score (10.8±2.3, range 2-23) was linked to two-year persistence (OR per point increment 1.29, $p<0.01$). The predictive properties of the TR-AD score appeared modest (AUC=0.66).

Limitations

In the current study, treatment history and ongoing treatments were retrospectively assessed. It was not evaluated whether prior treatments failed or succeeded.

Conclusions

The results in the current study suggest that when assessing TR-AD and designing a treatment plan, evaluations of treatment history should be accompanied with assessments of clinical characteristics. The dimensional TR-AD measurement presented here could be used for this purpose.

Introduction

Many patients with Anxiety Disorders experience suboptimal treatment results.[1-4] In clinical practice, patients are considered to have treatment resistant anxiety disorders (TR-AD) when evidence-based treatments do not yield sufficient symptom reduction, symptom severity is substantial and pseudo-resistance has been ruled out.[5-7] Pseudo-resistance may be due to unrecognized anxiogenic factors such as coffee or substances, or to inadequate treatment type, insufficient treatment duration, inadequate dosage regimes or nonadherence to treatments.[7,8] After adequate treatment, up till 30-60% of patients with anxiety disorders have substantial and impairing remaining symptoms.[1,9,10] Furthermore, even after successful treatments, around 12-20% of patients show a relapse after three years,[11] while relapse rates were approximately 50% in young patients after six years.[12] Patients with suboptimal treatment outcomes show increased levels of disability, comorbidity, loss of (work) functioning, reduced quality of life, increased mortality and higher health care costs.[13] Identifying TR-AD is thus of great clinical relevance due to its negative impact on subsequent course. However, the concept of TR-AD is poorly demarcated with many different definitions currently in use.[5,7] Generally, TR-AD is considered a dichotomous concept, i.e. patients either have treatment resistance or don't have treatment resistance. Using a dimensional approach to TR-AD might be more beneficial. Thereby, a sound understanding of TR-AD and a way to assess TR-AD in clinical care are much needed.

A dimensional approach for assessing the TR in depression (TR-D) exists in the Maudsley Staging Method (MSM), which is based on empirically validated aspects of TR-D.[14] The MSM and an adaptation, the Dutch Measure for quantification of Treatment Resistant Depression (DM-TRD) have proven useful tools in clinical care in depression.[15-17] Depressed patients with high levels of TR-D show a less favorable long term course in comparison to those with lower levels of TR-D.[18] In analogy to this, the current levels of TR-AD are assumed to be related to subsequent course in anxiety disorders in patients receiving treatments.[19,20]

From a recent systematic review into definitions for TR-AD various clinical aspects were identified as criteria for TR-AD.[5] These criteria for TR-AD included symptom severity, presence of functional impairments, psychiatric comorbidity, previous duration of symptoms, number of adequate pharmacological and psychological treatments. However, these aspects

of TR-AD have not been studied in concert in a sample of anxiety disorder patients who receive treatment. Furthermore, no dimensional measurement instrument exists for assessing levels of TR-AD, and hence, it is unknown whether such a quantified TR-AD measurement could be useful in clinical care.

The aim of this study is to operationalize a dimensional assessment of TR-AD. First, it will be assessed whether individual aspects of TR-AD derived from the literature are related to course during treatment in a sample with varying levels of treatment resistance at baseline. Next, by tailoring the DM-TRD for use in anxiety disorders, we will develop a dimensional tool to assess the level of TR-AD, expressed in a TR-AD score. The association between the TR-AD score and anxiety disorder status will be described in this sample of anxiety disorder patients who received treatments over the course of the follow-up period. Finally, the predictive properties of this newly developed dimensional TR-AD score will be examined.

Methods

Study sample

Subjects were derived from the ongoing Netherlands Study of Depression and Anxiety (NESDA), a naturalistic cohort study that examines the course of depression and anxiety in adults. At baseline, 2,981 subjects were included from the community, primary care and specialized mental health care. They had a major depressive disorder (MDD, *n*=1,222, 41%), anxiety disorder (*n*=1,305, 44%) or were healthy controls (*n*=632, 22%). Patients with a primary diagnosis of psychotic disorders, obsessive compulsive disorder, bipolar disorders, or severe substance abuse disorders were excluded. Baseline data were collected from 2004-2007 and two-year follow-up data from 2006-2009. Full methods for NESDA were previously described in detail.[21]

For the purpose of this study, we selected 1,305 subjects from all inclusion sites with a current anxiety disorder diagnosis at baseline: panic disorder (PD), agoraphobia, generalized anxiety disorder (GAD) and social anxiety disorder (SAD). By including subjects from all sampling sites we ensured inclusion of subjects with varying levels of TR-AD. Only subjects who at two-year follow-up reported to have received treatment in the period between baseline and follow-up were selected for the purpose of this study. The

reported treatments included psychotherapy, pharmacotherapy, or regular appointments with a mental health care worker. Ongoing treatments between baseline and follow-up were retrospectively assessed at follow-up using the Perceived Need for Care Questionnaire (PNCQ) [22]. By selecting subjects with confirmed ongoing treatments it was possible to assess the association of aspects of TR-AD with long term outcomes during treatment. Psychiatric disorders at baseline and two-year follow-up were assessed according to DSM-IV criteria with the Composite International Diagnostic Interview (CIDI, version 2.1).[23–25] The CIDI is a structured interview which was conducted by trained research staff. The CIDI was shown to have good overall reliability and validity and is frequently used worldwide.[24]

A total of 230 subjects who did not participate in the two-year follow-up were excluded, 378 subjects who did not receive treatment between baseline and follow-up, and a further 18 subjects with incomplete baseline data, yielding a total sample of 679 subjects. The 230 subjects who were excluded due to missing follow-up data did not differ significantly from included subjects in age, gender, type of anxiety disorder, presence of psychiatric comorbidity, and various clinical characteristics (p>0.10). However, they had fewer education years (11.1 vs 12.0, p<0.001).

Aspects of TR-AD

A recent systematic review into TR-AD identified various criteria for TR-AD. These included minimal number of adequate pharmacological treatments, minimal number of adequate psychological treatments, minimal anxiety disorder symptom severity, presence of functional impairments, presence of psychiatric comorbidity, and minimal previous duration of symptoms.[5] These literature-derived criteria were included in the current study as aspects of TR-AD and were assessed at baseline. Treatments were assessed twice. The first assessment was done retrospectively at baseline with regard to the three years prior to baseline and the second assessment was done at follow-up with regard to the period between baseline and follow-up measurements.

First- and second line pharmacotherapy

For the purpose of the current study, 'adequate pharmacologic treatment' was defined as first- and second-line evidence-based anti-anxiety drugs taken daily for at least two months at an effective dosage. All current pharmacotherapeutic use was assessed using inspection of medication containers and coded according to Anatomical Therapeutic Chemical (ATC) codes at baseline.[26] Historic use of anti-anxiety drugs during the

three-year period prior to baseline was assessed using retrospective self-reports. Data from current and previous three year pharmacotherapeutic trials were combined into number of first-line and number of second-line pharmacotherapy trials. Categorization into first-line and second-line treatments was based on National Institution Clinical Excellence anxiety disorder guidelines and the Dutch anxiety disorder treatment guidelines.[27–29] SSRIs and SNRIs (venlafaxine or duloxetine) were considered first-line anti-anxiety pharmacotherapy trials.[27–29] Second-line anti-anxiety pharmacotherapy trials included tricyclic antidepressants, tetracyclic antidepressants (mirtazapine and trazodone), monoamine-oxidase inhibitors, high potency benzodiazepines (alprazolam, clonazepam, lorazepam, diazepam and bromazepam), pregabaline and buspirone.[27–29] For each drug the daily dosages subjects reported were checked and were required to meet the registered daily derived dosage (DDD) for that drug.[26] In subjects who did not know the quantity of previous medication, we estimated their daily dosage using the lowest available quantity tablets available and multiplying that with the number of tablets they reported taking daily.

Adequate psychological treatments

Psychological treatments prior to baseline assessment were assessed using the PNCQ, in which subjects were asked which types of care they received.[22] From these, we included psychotherapy trials when subjects reported having had at least ten sessions with a psychologist, individual psychiatrist or psychotherapist, or when they reported at least 16 sessions at a mental health care institution.

Anxiety disorder symptom severity

Symptom severity was defined using both severity of anxiety (according to the Beck Anxiety Inventory, BAI)[30] and severity of avoidance behaviors (according to the Fear Questionnaire, FQ).[31] The BAI is a 21-item self-report questionnaire that measures severity of anxiety. The BAI has adequate psychometric properties.[32,33] Anxiety severity was categorized according to previously identified cut-off values: mild (BAI < 10), moderate (BAI 10 < 30) and severe (BAI ≥ 30).[33,34]

The FQ is a 15-item questionnaire that measures avoidance in three domains: agoraphobia (Ag), social phobia (So) and blood injury phobia. The latter domain was omitted for the purpose of this study as specific phobias fall outside the scope of this study. Different cut-off values are used for the FQ subscales. Cut-off values for clinically relevant levels of agoraphobic

avoidance (FQ Ag) vary from ≥15 to ≥19, whereas cut-off values for clinically relevant social phobia avoidance (FQ So) vary from ≥12 to ≥18.[35,36] The severity of avoidance behaviours was categorized according to FQ scores: mild severity was defined as FQ Ag below 15 and FQ So below 12, moderate severity was defined as FQ Ag between 15 and 19 and FQ So between 12 and 18, severe was defined as FQ Ag at 19 or above and FQ So at 18 or above. For the purpose of data analysis, for each patient the symptom severity was defined by the highest value on either the BAI or the FQ according to the cut-off values described, creating a single variable for symptom severity (mild, moderate or severe), as done before.[34]

Presence of functional impairments

Levels of functional impairment were assessed using the WHO-DAS, 32 item version.[37] The WHO-DAS measures disability in six domains: cognition, mobility, self-care, interpersonal interactions, household activities, and participation in society on a 5-point Likert scale with item-scores ranging from 0 (no difficulties) to 4 (extreme difficulties/cannot do). A total disability score was calculated by adding all 32 item-scores. The WHO-DAS has excellent internal consistency (α=0.95).[37] Total scores were compared against a conversion table to derive the general population percentile score for each subject.[38] Levels of functional impairments were categorized into 'low' (< 50th general population percentile), 'moderate' (50th < 75th general population percentile) and 'severe' (> 75th general population percentile).

Presence of psychiatric comorbidity

The presence of current (6-month) dysthymia, major depressive disorder and alcohol dependency were assessed with the CIDI and psychiatric comorbidity was considered present if any of these disorders were classified.

Previous duration of symptoms

Previous duration of symptoms was retrospectively assessed at baseline using the Life chart method.[39] It uses a calendar based approach to provide memory anchors for subjects in order to assess the presence of anxiety and avoidance symptoms during each month of the current and four previous calendar years. The LCI has adequate reliability and validity.[40] If either anxiety or avoidance was present in at least 80% of the months prior to baseline, previous duration was 'long'. If anxiety or avoidance was present in 30-80% of the time, previous duration was 'intermediate', and if anxiety or avoidance was present in less than 30% of the previous months, previous duration was 'short'.

Dimensional TR-AD score

Levels of TR-AD were assessed using a combined score comprised of the different individual aspects of TR-AD. Scoring for this dimensional score was based on the Dutch Measure for quantification of Treatment Resistance in Depression (DM-TRD).[17] The scoring system used in the DM-TRD was maintained as aspects of TR-AD were identical to aspects for TR-D used in the DM-TRD (see Panel 1 for scoring). Each of the aspects of TR-AD were scored in accordance with the DM-TRD to derive a single score for each subject. This yielded a dimensional measurement instrument for levels of TR-AD with a potential range of 2-23.

Main outcome variable

Persistence of anxiety disorders was assessed at two-year follow-up, defined as 6-month presence of PD, GAD, agoraphobia or SAD, diagnosed with the CIDI.

Main analyses

In order to relate baseline individual aspects of TR-AD to outcome at two-year follow-up, bivariate logistic regression analyses were performed. Odds Ratios (OR) for each aspect of TR-AD were reported. Next, logistic regression analyses were performed to link the dimensional TR-AD score assessed at baseline to outcome at two-year follow-up. To assess the predictive properties for the dimensional TR-AD score, the Youden-index,[41] which is indicative of the most optimal cut-off score; sensitivity, specificity, positive predictive values and negative predictive values for different cut-off values were calculated. A Receiver Operator Curve (ROC) was plotted and area under the curve (AUC) was calculated. Statistical analyses were performed using IBM SPSS 23.

Sensitivity analyses

In the design of this study it was retrospectively assessed at follow-up whether subjects received treatment in the period leading up to follow-up. In this approach it was not clear, however, when treatments were initiated. Therefore, a sensitivity analysis was performed in which the main analyses were repeated in a subsample who presented at specialized mental health care at baseline (n=405). Within the design of NESDA, all subjects that were included at specialized mental health care institutions were initiated evidenced-based care directly after inclusion into the study. This yielded a sample in which treatments were initiated at the moment of inclusion and in which treatments were ongoing at follow-up.

Panel 1. Dimensional assessment of degree of Treatment Resistant anxiety disorders (TR-AD).

Item and specification	score
Symptom severity [1]	
Mild	1
Moderate	3
Severe	5
Functional impairments [2]	
None	0
Low	1
Moderate	2
Severe	3
Psychiatric comorbidity [3]	
No	0
Yes	2
Previous duration of anxiety [4]	
Short	1
Intermediate	2
Long	3
First-line antianxiety pharmacotherapy trials [5]	
0	0
1-2	1
3-4	2
5-6	3
7-10	4
>10	5
Second-line antianxiety pharmacotherapy trials [6]	
0	0
1-2	1
3-4	2
5-6	3
Psychotherapy trials [7]	
0	0
1	1
≥2	2
(Degree of TR-AD)	(2-23)

[1] *mild*: Beck Anxiety Inventory (BAI) <10, Fear Questionnaire (FQ), agoraphobia subscale (Ag) <15 and FQ, social phobia (So) subscale <12; *moderate*: BAI 10<30, or FQ (Ag) 15<19, or FQ (So) 12<18; *severe*: BAI ≥30, FQ (Ag) ≥19, or FQ (So) ≥18).
[2] World Health Organisation Disability Assessment Schedule 2.0 general population percentile scores: none<25th, low: 25th-50th, moderate: 50th-75th, severe >75th
[3] major depressive disorder, dysthymia, alcohol dependency.
[4] *short*: <30% of months during the previous five calendar years spent with symptoms, *intermediate*: 30<80% of months during the previous five calendar years spent with symptoms, *long*: >80% of months during the previous five calendar years spent with symptoms.
[5] selective serotonin reuptake inhibitors (SSRIs), serotonin-norepinephrine reuptake inhibitors (SNRIs).
[6] tricyclic antidepressants (TCAs), monoamine oxidase inhibitors (MAO-i), tetracyclic antidepressants, high potency benzodiazepines, pregabalin, and buspirone.
[7] at least 10 consultations with a psychologist, psychotherapist, or psychiatrist, or at least 16 consultations at a mental health center

Results

Sample

The sample of *n*=679 subjects was 40.4 ± 11.1 years old, 452 subjects were female (66.6%) and subjects received 12.0 ± 3.3 years of education. The sample consisted of subjects diagnosed with PD (*n*=348, 51.3%), agoraphobia (*n*=84, 12.4%), SAD (*n*=373, 54.9%) and GAD (*n*=270, 39.8%). Prior to inclusion, subjects received 0.44 ± 0.65 first-line pharmacotherapy treatments, 0.12 ± 0.38 second-line pharmacotherapy treatments and 0.14 ± 0.36 psychotherapy treatments. The total number of previous treatments (pharmacotherapy or psychotherapy) was 0.69 ± 0.94, with a range of 0-5 previous treatments. The mean TR-AD-score was 10.8 ± 2.3. See Table 1 for baseline characteristics.

Table 1. Baseline characteristics of treated anxiety disorder sample (*n*=679).

	n (%)	mean ± SD
Demographic characteristics		
Age		40.4 ± 11.9
Female gender	452 (66.6%)	
Education		12.0 ± 3.3
Basic	47 (6.9%)	
Intermediate	426 (62.7%)	
High	206 (30.3%)	
Sampling site		
General population	49 (7.2%)	
Primary care	225 (33.1%)	
Specialised mental health care	405 (59.6%)	
Type of anxiety disorder		
panic disorder	348 (51.3%)	
agoraphobia	84 (12.4%)	
social anxiety disorder	373 (54.9%)	
generalized anxiety disorder	270 (39.8%)	
Aspects of Treatment Resistance		
Number of first-line pharmacotherapy trials		0.44 ± 0.65
0	433 (63.8%)	
1	208 (30.6%)	
2	31 (4.6%)	
3 or more	7 (1.0%)	

Table continues

Table 1. Continued

Number of second-line pharmacotherapy trials			0.12 ± 0.38
0		615 (90.6%)	
1		52 (7.7%)	
2 or more		12 (1.8%)	
Number of psychotherapy trials			0.14 ± 0.36
0		591 (87.0%)	
1		84 (12.4%)	
2		4 (0.6%)	
Symptom severity			
Mild		52 (7.7%)	
Moderate		272 (40.1%)	
Severe		355 (52.3%)	
Levels of functional impairments			
Low	<50th percentile of general population	4 (0.6%)	
Moderate	50th-75th percentile	39 (5.7%)	
Severe	>75th percentile	636 (93.7%)	
Psychiatric comorbidity			
No		213 (31.4%)	
Yes		466 (68.6%)	
Previous duration of anxiety (previous five years)			
Short	<30% months	229 (33.7%)	
Intermediate	30-80% months	197 (29.0%)	
Long	>80% months	253 (37.3%)	
TR-AD score (2-23)			10.8 ± 2.3 (range 2-16)

Main analyses

Most baseline aspects of TR-AD were significantly associated with persistence of anxiety disorders at two-year follow-up in bivariate models (See Table 2). For example, high levels of functional impairments were related to persistence: OR=2.90; 95% Confidence Interval (CI), 1.51-5.63, p=0.001. Symptom severity was strongly associated with persistence (OR severe symptoms vs mild symptoms =6.48; 95% CI, 3.29-12.8, p<0.001). Surprisingly, previous pharmacotherapy treatments and previous psychotherapy treatments were not significantly related to two-year persistence, nor was the combined total number of treatments. Higher scores on the dimensional TR-AD measurement were associated with

two-year persistence: OR (per point increment) =1.29; 95% CI, 1.20-1.39, p<0.001, which translates into OR (per SD increase) =1.81; 95%CI, 1.53-2.15, p<0.001.

Table 2. Bivariate associations between baseline aspects of treatment-resistance in anxiety disorders with two-year persistence in anxiety disorder patients (n=679).

Bivariate baseline predictors		Two-year persistence OR (95% CI)
Age		1.01 (0.99-1.02)
Female gender		0.87 (0.63-1.20)
Education years		**0.95 (0.90-0.99)**
Symptom severity		
	Mild	1.00 (ref)
	Moderate	3.68 (1.85-7.30)
	Severe	**6.48 (3.29-12.8)**
High levels of functional impairment		**2.90 (1.51-5.63)**
Presence of psychiatric comorbidity		**1.73 (1.25-2.39)**
Previous duration of anxiety (% of months during last 5 years)		
	Short (<30%)	1.00 (ref)
	Intermediate (30-80%)	**2.22 (1.49-3.21)**
	Long (>80%)	**2.79 (1.94-4.03)**
Number of first-line pharmacotherapy trials		1.10 (0.87-1.39)
Number of second-line pharmacotherapy trials		1.39 (0.91-2.12)
Number of psychotherapy trials		1.11 (0.73-1.69)
Total number of treatments		1.12 (0.95-1.32)
Degree of TR-AD		**1.29 (1.20-1.39)**

Boldface indicates p< 0.05

The optimal cut-off for the dimensional TR-AD score based on the highest Youden index was 11 or higher. When using this cut-off value, sensitivity was 0.70 and specificity 0.57. Positive predictive value (PPV) and negative predictive value (NPV) were modest: 0.68 (PPV) and 0.60 (NPV). See figure 1 for the ROC using baseline levels of TR-AD as predictor for two-year persistence. From this ROC, the AUC was calculated at 0.66 (see Figure 1).

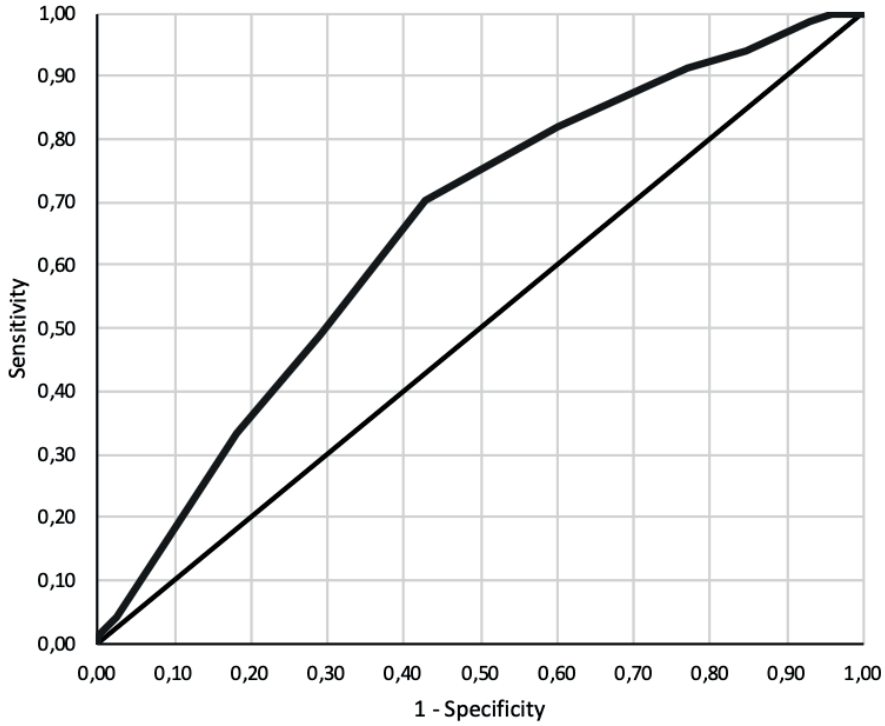


Figure 1. Receiver operator curve and predictive properties for baseline degree of TR-AD score on two-year anxiety disorder persistence.

Sensitivity analyses

The main analyses were repeated in a subsample of patients who were included in specialized mental health care (n=405). In comparison with subjects that were sampled elsewhere (n=274), this subset was younger (mean age= 38.3 versus 43.6, p<0.001), included a lower proportion of female patients (62.7% vs. 72.3%, p=0.006), higher proportions of PD diagnoses (56.5% vs. 43.4%, p=0.001), higher proportions of GAD diagnoses (42.7% vs. 35.4%, p=0.03), higher symptom severity (56.5% in severe severity vs. 46.0%, p=0.03), higher proportions of psychiatric comorbidity (75.3% vs. 58.8%, p<0.001), longer duration of symptoms (31.6% with moderate duration vs. 25.2%, p=0.04), more often had previous first-line pharmacological treatments (mean number=0.54 vs. 0.28, p<0.001), more often had previous second-line pharmacological treatments (mean number= 0.16 vs. 0.04, p<0.001). The subsample had similar educational levels (mean education years= 11.8 versus 12.2, p=0.15), similar proportions of

social phobia diagnoses (57.0% vs. 51.8%, p=0.10), similar overall levels of functioning (92.8% in fourth WHODAS quartile vs. 94.9%, p=0.53) and similar number of psychological treatments (mean number of treatment= 0.15 vs. 0.12, p=0.27) in comparison with the subjects sampled in other inclusion sites. These differences in baseline characteristics were to be expected, as in the Dutch health care system more severely affected patients are more likely to be referred to specialized mental health care. In spite of these baseline differences, all analyses showed comparable results to the whole sample. The same predictors for course after treatment were found in this subsample (see Table 3) and the psychometric properties for the dimensional TR-AD score were very similar (see Figure 2).

Table 3. Sensitivity analysis: Two-year bivariate associations between clinical characteristics and previous treatment types with persistence in anxiety disorder patients who were initiated treatments in specialized mental health care (n=405) with confirmed ongoing treatment after two years.

		Two-year persistence
Bivariate baseline predictors		OR (95% CI)
Age		1.01 (0.99-1.02)
Female gender		0.85 (0.57-1.29)
Education years		**0.93 (0.88-0.99)**
Symptom severity		
	Mild	1.00 (ref)
	Moderate	**4.84 (1.86-12.6)**
	Severe	**7.14 (2.79-18.3)**
High levels of functional impairment		**2.90 (1.31-6.40)**
Presence of psychiatric comorbidity		**1.86 (1.18-2.94)**
Previous duration of anxiety (% of months during last 5 years)		
	Short (<30%)	1.00 (ref)
	Intermediate (30-80%)	**2.66 (1.59-4.43)**
	Long (>80%)	**3.22 (1.96-5.29)**
Number of first-line pharmacotherapy trials		1.06 (0.80-1.41)
Number of second-line pharmacotherapy trials		1.16 (0.72-1.85)
Number of psychotherapy trials		1.07 (0.63-1.81)
Total number of treatments		1.07 (0.88-1.29)
Degree of TR-AD		**1.29 (1.17-1.42)**

Boldface indicates p<0.05

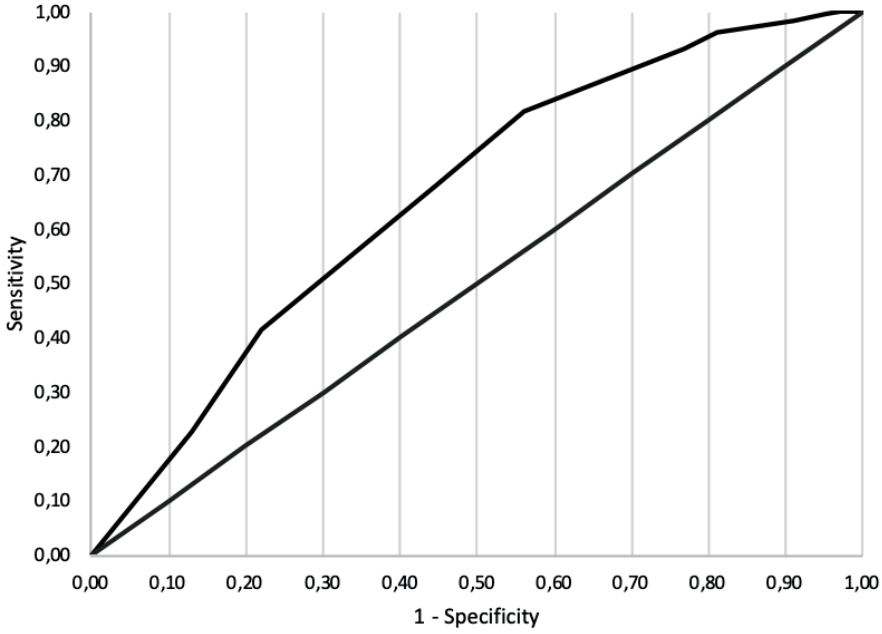


Figure 2. Sensitivity analysis: Receiver operator curve and predictive properties for baseline degree of TR-AD score on two-year anxiety disorder persistence in sample that initiated treatments after inclusion into the cohort (*n*=405).

Discussion

Despite the relevance in clinical care, defining and assessing treatment resistance in anxiety disorders (TR-AD) has hardly been a focus in scientific research. This study examined aspects of TR-AD and examined whether combining these aspects into a dimensional TR-AD score could adequately predict two-year course in a sample of anxiety disorder patients who received treatment during a two-year period.

The first aim was to empirically validate literature-derived aspects of TR-AD. Higher baseline symptom severity, levels of functional impairment, presence of psychiatric comorbidity and longer previous duration of symptoms were critically related to persistence of anxiety disorders at two-year follow-up in a treated sample. These clinical characteristics are sometimes identified as prognostic factors in clinical care, and are included as such in treatment guidelines.[42] Surprisingly, in our study treatment history showed no significant impact on two-year course. This is a counterintuitive

finding because failed treatment is regarded the core criterion for TR in any disorder and hence, assessment of treatment history is common practice in clinical care.[2] Additionally, treatment history is the main criterium on which treatment guidelines base their recommendations for evidence-based stepped-care algorithms in which nonresponse on a low intensity treatment usually leads to recommending more aggressive treatments.[28,42] Therefore, in current stepped-care algorithms the number of failed treatments indicates the current level of TR-AD. In our study, it was unclear whether previous treatments had failed or were successful at the time, while only failed treatments define TR. Moreover, a previous study showed that adherence to treatment guidelines was suboptimal in the NESDA sample. In one third of included patients, the treatments provided were fully adherent to treatment guidelines, while in a small majority, the treatments received were not fully adherent to treatment guidelines.[43] Therefore, the method of data collection could have contributed to the lack of association between treatment history and two-year course. However, a recent study into TR factors in depressed NESDA subjects used the same approach to assess treatment history and this study did show an association between treatment history and two-year outcomes in depression.[18] In depression this association is in line with the existing literature, for example, the STAR*D trial empirically showed that a history of pharmacotherapy trial failures preceded subsequent lack of treatment response.[44] Even as this mechanism was never empirically demonstrated in anxiety disorders, it is generally assumed to be present as well.[19,20] It is possible that the association between treatment history and subsequent course during treatment in anxiety disorders is less robust in comparison with depression. This could be due to differences in naturalistic course: anxiety disorders more often show chronicity, whereas MDD more often shows an episodic course.[45] The results in this study could be indicative of the shortcomings of using treatment history as the cornerstone in assessments of TR-AD and in stepped-care treatment algorithms and it is unclear whether this cornerstone position of assessing treatment history is fully warranted. Findings of the present study suggest that higher levels of anxiety duration, symptom severity, psychiatric comorbidity and functional impairments contribute to higher levels of TR-AD. Based on the findings presented here, these aspects should be used alongside assessments of previous treatment failures in the process of assessing treatment resistance and designing treatment plans.

A second aim of the present study was to develop a dimensional tool to assess the levels of TR-AD, reflected as a TR-AD score. We tailored the Dutch Measure for quantification of Treatment Resistance in Depression (DM-TRD) for use in anxiety disorders, based on literature-derived aspects of TR-AD. Each point increment in the dimensional TR-AD score was associated with increased Odds of 1.29 on persistence of anxiety disorders, which translates into increased Odds of 1.81 per SD increase. Predictive power was moderate, with an AUC of 0.66. This magnitude of effect similar is to that in TR depression (TR-D). The level of TR-D, as measured with the MSM, was linked to "persistent depression" (>50% of the time spent with depressive symptoms) at two-year follow-up: each point increment was associated with 1.40 increased Odds for persistent depression.[18] But whereas the dimensional TR-AD score presented here had a range of 2-23 with a sample SD of 2.34, the MSM has a range of 3-15 with a sample SD of 1.22 [18]. Therefore, the Odds increase per SD increment of 1.81 are higher in the current TR-AD sample in comparison with the TR-D sample. Comparing the AUCs was not possible as no AUC was reported in the paper by van Belkum et al (2018).[18] Thereby, the current dimensional TR-AD score shows promise as a measurement tool in TR-AD. It could be used in the process of assessing treatment resistance and when designing a treatment plan.

Strengths of the current study include the use of literature-derived aspects of TR-AD, the longitudinal approach and applying strict requirements to treatment regimens to take pseudoresistance into account. Some limitations should also be noted. First, we aimed to study effects of aspects of TR-AD on anxiety course in subjects receiving treatments. This approach was chosen as it would enable identifying patients with decreased chances of beneficial treatment effects, which is clinically very relevant. In our main sample, for some patients it was unclear when these treatments were initiated. This could have led to reduced external validity for our results to be interpreted in real world treatment-seeking samples. When we repeated the main analyses in a subsample who initiated treatments directly after inclusion into the cohort (n=405) the same results after two-year follow-up were found. Therefore, it was concluded that the uncertainty with regard to the moment of initiation of treatment was not a confounding factor in this study. The results from the sensitivity analyses suggest that the same underlying processes are present in both subsamples. Second, some limitations with regard to the assessments of previous treatments existed. The present study relied on current and previous three-year treatment history as information on lifetime treatment history was not available. As

a result, impact of previous treatments in assessing TR-AD may have been underestimated, as *lifetime* treatment history is considered important in TR.[2,19,20,46,47] Also, a number of subjects reported having no previous treatments. This could have led to a bias towards the null hypothesis if these subjects had failed treatments before the assessment period (over three years ago). Furthermore, the current design in which previous treatments were assessed retrospectively via self-report introduced the risk of a recall bias. Patients with anxiety disorders are somewhat prone to memory biases, especially if the memories are emotionally laden.[48] A recall bias would lead to underreporting of previous treatments, which also leads to a bias towards the null hypothesis. Additionally, as mentioned previously, no assessments of previous treatment effects were undertaken. This could also have led to a bias towards the null hypothesis as only failed treatments are considered to be of importance in assessments of TR. Moreover, evidence-based psychotherapies for anxiety disorders include homework assignments and incorporation of exposure interventions. These aspects of psychotherapies are likely related to treatment outcomes in anxiety disorders.[49] In the design of this study, the exact contents of the psychotherapies were not known. As a result of these limitations, the lack of association between treatment history and subsequent course in anxiety disorder patients should be replicated in a different cohort in which these shortcomings in assessments of previous treatments are not present. Finally, some disorders with high levels of comorbidity with anxiety disorders fell outside the scope of NESDA. For instance, specific phobias and obsessive compulsive disorder were not routinely included. Likely, presence of these disorders has impact on the treatment history and course in anxiety. Therefore, the current findings cannot be generalized to populations with these comorbidities.

Future studies could aim to replicate the associations between a dimensional TR-AD score, as measured with the measurement instrument presented in this paper, with treatment effects in anxiety disorder patients. For future studies using a retrospective design it might be beneficial to involve pharmacists in determining previous pharmacological treatments. Moreover, asking patients to specify certain aspects of previous psychotherapies might increase identification of adequate psychotherapeutic treatments. For instance, in addition to asking the number of sessions it might be beneficial to ask whether homework assignments were given and whether the treatments included exposure interventions. Also, further research into the role of failed anxiety disorder treatments with regard to course during subsequent treatments is warranted. Ideally, a randomized trial

investigating a stepped-care algorithm in anxiety disorders, like STAR*D in depression, should be performed. This would reduce the risks of bias due to underreporting of previous treatments as the treatments would be assessed prospectively. This could further uncover the role of different aspects of TR-AD during subsequent treatments and improve treatment decision making within a stepped-care algorithm.

In summary, we showed associations between several literature-derived aspects of TR-AD with persistence of anxiety disorders at two-year follow-up in a sample receiving treatment and developed a dimensional tool to assess TR-AD that showed promise. There was a clear association between the score on this TR-AD measurement with persistence of anxiety disorders after a two-year follow-up period during which respondents received treatments. The association of TR-AD with course during treatment seems more driven by baseline clinical characteristics in comparison to treatment history. The lack of significant associations between treatment history and course during treatment demonstrated in this study warrants further investigation into the role of previous failed treatments in anxiety disorder clinical care. Ideally, all aspects of TR-AD should be investigated prospectively in a cohort of anxiety disorder patients receiving evidence-based treatments to determine which of these factors should be most central in treatment decision making in anxiety disorder patients. The current research suggests that assessments of severity, duration, disability and psychiatric comorbidities could have the highest contribution in treatment decision making.

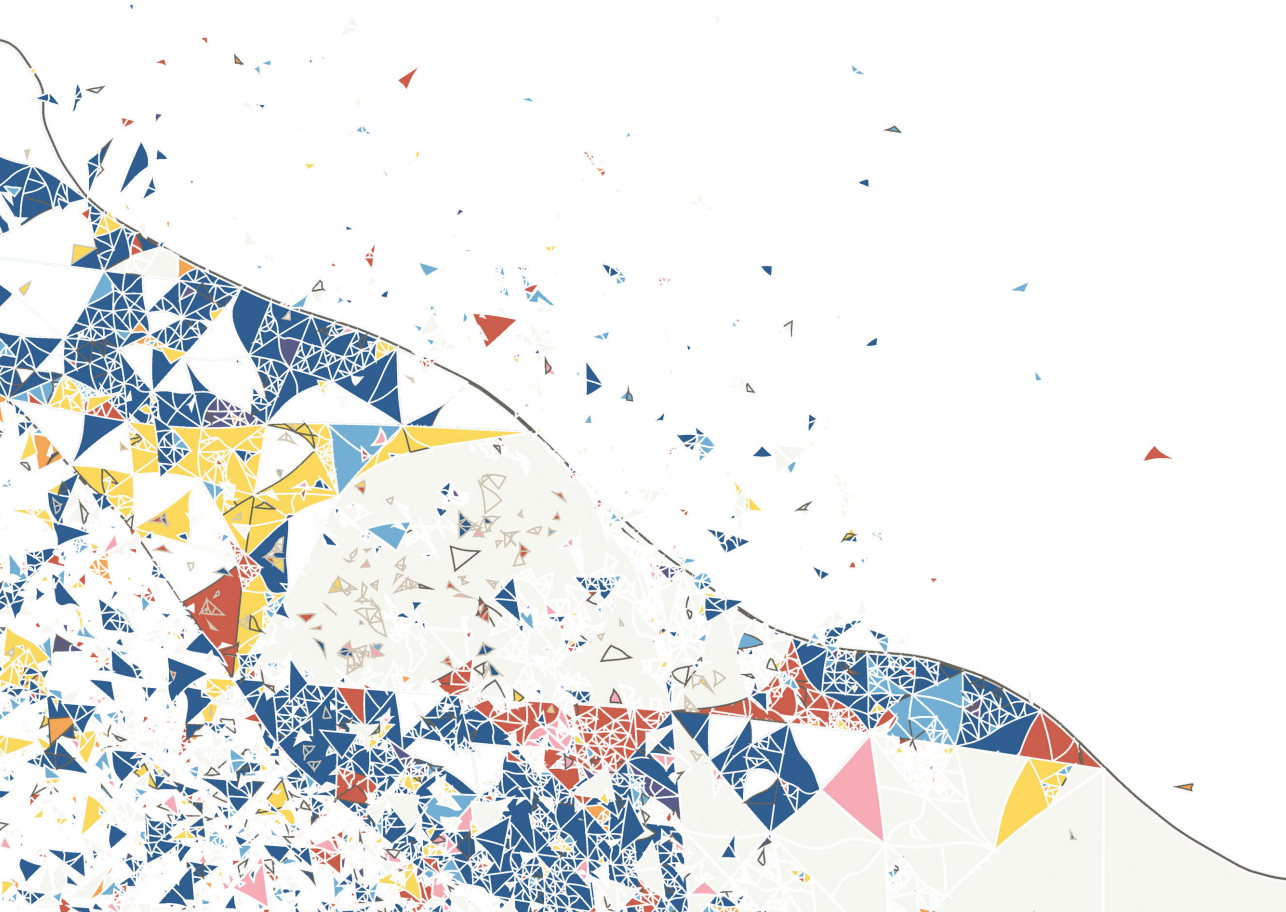
References

1. Bruce SE, Yonkers KA, Otto MW, et al. Influence of Psychiatric Comorbidity on Recovery and Recurrence in Generalized Anxiety Disorder, Social Phobia, and Panic Disorder: A 12-Year Prospective Study. *Am J Psychiatry*. 2005;162(6):1179-1187. doi:10.1176/appi.ajp.162.6.1179
2. Bystritsky A. Treatment-resistant anxiety disorders. *Mol Psychiatry*. 2006;11(9):805-814. doi:10.1038/sj.mp.4001852
3. Huh J, Goebert D, Takeshita J, Lu BY, Kang M. Treatment of Generalized Anxiety Disorder: A Comprehensive Review of the Literature for Psychopharmacologic Alternatives to Newer Antidepressants and Benzodiazepines. *Prim Care Companion CNS Disord*. 2011;13(2):e1-e9. doi:10.4088/PCC.08r00709
4. Ramsawh HJ, Raffa SD, Edelen MO, Rende R, Keller MB. Anxiety in middle adulthood: effects of age and time on the 14-year course of panic disorder, social phobia and generalized anxiety disorder. *Psychol Med*. 2009;39(4):615-624. doi:10.1017/S0033291708003954
5. Bokma WA, Wetzer GAAM, Gehrels JB, Penninx BWJH, Batelaan NM, van Balkom ALJM. Aligning the many definitions of treatment resistance in Anxiety Disorders: a systematic review. *Depress Anxiety*. 2019;36(9):801-812. doi:10.1002/da.22895
6. Fava GA, Rafanelli C, Tomba E. The Clinical Process in Psychiatry: A Clinimetric Approach. *J Clin Psychiatry*. 2012;73(2):177-184. doi:10.4088/JCP.10r06444
7. Roy-Byrne P. Treatment-refractory anxiety; definition, risk factors, and treatment challenges. *Dialogues Clin Neurosci*. 2015;17(2):191-206.
8. Farach FJ, Pruitt LD, Jun JJ, Jerud AB, Zoellner LA, Roy-Byrne PP. Pharmacological treatment of anxiety disorders: Current treatments and future directions. *J Anxiety Disord*. 2012;26(8):833-843. doi:10.1016/j.janxdis.2012.07.009
9. Tyrer P, Seivewright H, Johnson T. The Nottingham Study of Neurotic Disorder: Predictors of 12-year outcome of dysthymic, panic and generalized anxiety disorder. *Psychol Med*. 2004;34(8):1385-1394. doi:10.1017/S0033291704002569
10. Bandelow B, Zohar J, Hollander E, Kasper S, Möller H-J, WFSBP Task force on treatment guidelines for anxiety obsessive-compulsive post-traumatic stress disorders. World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for the pharmacological treatment of anxiety, obsessive-compulsive and post-traumatic stress disorders - First revision. *World J Biol Psychiatry*. 2008;9(4):248-312. doi:10.1080/15622970802465807
11. Nay W, Brown R, Roberson-Nay R. Longitudinal course of panic disorder with and without agoraphobia using the national epidemiologic survey on alcohol and related conditions (NESARC). *Psychiatry Res*. 2013;208(1):54-61. doi:10.1016/j.psychres.2013.03.006

12. Ginsburg GS, Becker EM, Keeton CP, et al. Naturalistic follow-up of youths treated for pediatric anxiety disorders. *JAMA psychiatry*. 2014;71(3):310-318. doi:10.1001/jamapsychiatry.2013.4186
13. Baxter AJ, Vos T, Scott KM, Ferrari AJ, Whiteford HA. The global burden of anxiety disorders in 2010. *Psychol Med*. January 2014:1-12. doi:10.1017/S0033291713003243
14. Fekadu A, Wooderson S, Donaldson C, et al. A multidimensional tool to quantify treatment resistance in depression: The Maudsley staging method. *J Clin Psychiatry*. 2009;70(2):177-184. doi:10.4088/JCP.08m04309
15. Fekadu A, Donocik JG, Cleare AJ. Standardisation framework for the Maudsley staging method for treatment resistance in depression. *BMC Psychiatry*. 2018;18(1):1-13. doi:10.1186/s12888-018-1679-x
16. van Dijk DA, van den Boogaard TM, Deen ML, Spijker J, Ruhé HG, Peeters FPML. Predicting clinical course in major depressive disorder: The association between DM-TRD score and symptom severity over time in 1115 outpatients. *Depress Anxiety*. 2018;(November):1-8. doi:10.1002/da.22865
17. Peeters FPML, Ruhe HG, Wichers M, et al. The Dutch Measure for quantification of Treatment Resistance in Depression (DM-TRD): an extension of the Maudsley Staging Method. *J Affect Disord*. 2016;205:365-371. doi:10.1016/j.jad.2016.08.019
18. van Belkum SM, Geugies H, Lysen TS, et al. Validity of the Maudsley Staging Method in Predicting Treatment-Resistant Depression Outcome Using the Netherlands Study of Depression and Anxiety. *J Clin Psychiatry*. 2018;79(1). doi:10.1103/PhysRevLett.113.113903
19. Ipser JC, Carey P, Dhansay Y, Fakier N, Seedat S, Stein DJ. Pharmacotherapy augmentation strategies in treatment-resistant anxiety disorders (Review). *Cochrane Database Syst Rev*. 2006;(4). doi:10.1002/14651858.CD005473.pub2
20. Cosci F, Fava GA. Staging of mental disorders: systematic review. *Psychother Psychosom*. 2013;82(1):20-34. doi:10.1159/000342243
21. Penninx BWJH, Beekman ATF, Smit JH, et al. The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *Int J Methods Psychiatr Res*. 2008;17(3):121-140. doi:10.1002/mpr.256
22. Meadows G, Harvey C, Fossey E, Burgess P. Assessing perceived need for mental health care in a community survey: development of the Perceived Need for Care Questionnaire (PNCQ). *Soc Psychiatry Psychiatr Epidemiol*. 2000;35(9):427-435. doi:10.1007/s001270050260
23. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR Fourth Edition (Text Revision)*.; 2000.
24. Wittchen H. Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): a critical review. *J Psychiatr Res*. 1994;28:57-84.

25. World Health Organization. *World Health Organization, Composite International Diagnostic Interview (CIDI), Core Version 2.1*. Geneva; 1998.
26. WHO Collaborating Centre for Drug Statistics Methodolog. ATC/DDD Index 2019. https://www.whocc.no/atc_ddd_index/. Published 2019.
27. National Institute for Health and Clinical Excellence. *Generalised Anxiety Disorder and Panic Disorder in Adults: Management*. Manchester, United Kingdom; 2011.
28. National Institute for Health and Clinical Excellence. Social Anxiety Disorder: recognition, assessment and treatment. 2013;(May 2013). nice.org.uk/guidance/cg159.
29. Trimbos instituut. *Multidisciplinaire Richtlijn Angststoornissen (3e Revisie, 2013)*.; 2013.
30. Beck AT, Steer RA, Carbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clin Psychol Rev*. 1988;8(1):77-100. doi:10.1016/0272-7358(88)90050-5
31. Marks IM, Mathews AM. Brief standard self-rating for phobic patients. *Behav Res Ther*. 1979;17(3):263-267. doi:10.1016/0005-7967(79)90041-X
32. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: Psychometric properties. *J Consult Clin Psychol*. 1988;56:893-897.
33. Kabacoff RI, Segal DL, Hersen M, Van Hasselt VB. Psychometric properties and diagnostic utility of the Beck Anxiety Inventory and the State-Trait Anxiety Inventory with older adult psychiatric outpatients. *J Anxiety Disord*. 1997;11(1):33-47. doi:10.1016/S0887-6185(96)00033-3
34. Bokma WA, Batelaan NM, Hoogendoorn AW, Penninx BWJH, van Balkom AJLM. A clinical staging approach to improving diagnostics in anxiety disorders: Is it the way to go? *Aust New Zeal J Psychiatry*. 2020;54(2):173-184.
35. Schadé A, Marquenie LA, van Balkom AJLM, Koeter MWJ, Van Den Brink W, Van Dyck R. Effectiviteit van een toegevoegde behandeling voor angstklachten bij alcoholafhankelijke patienten met een fobische stoornis. *Tijdschr Psychiatr*. 2008;50(3):137-148.
36. Van Zuuren FJ. The Fear Questionnaire. Some data on validity, reliability and layout. *Br J Psychiatry*. 1988;153(NOV.):659-662. doi:10.1192/bjp.153.5.659
37. Chwastiak LA, Von Korff M. Disability in depression and back pain: Evaluation of the World Health Organization Disability Assessment Schedule (WHO DAS II) in a primary care setting. *J Clin Epidemiol*. 2003;56(6):507-514. doi:10.1016/S0895-4356(03)00051-9
38. Üstün T, Kostanjsek N, Chatterji S, Rehm J. *Measuring Health and Disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0)*.; 2010.
39. Lyketsos CG, Nestadt G, Cwi J, Heithoff K, et al. The Life Chart Interview: A standardized method to describe the course of psychopathology. *Int J Methods Psychiatr Res*. 1994;4(3):143-155.

40. Warshaw MG, Keller MB, Stout RL. Reliability and validity of the longitudinal interval follow-up evaluation for assessing outcome of anxiety disorders. *J Psychiatr Res*. 1994;28(6):531-545. doi:10.1016/0022-3956(94)90043-4
41. Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. *Biometrical J*. 2005;47(4):458-472. doi:10.1002/bimj.200410135
42. Baldwin DS, Anderson IM, Nutt DJ, et al. Evidence-based pharmacological treatment of anxiety disorders, post-traumatic stress disorder and obsessive-compulsive disorder: a revision of the 2005 guidelines from the British Association for Psychopharmacology. *J Psychopharmacol*. 2014;28(5):403-439. doi:10.1177/0269881114525674
43. van Dijk MK, Oosterbaan DB, Verbraak MJPM, Hoogendoorn AW, Penninx BWJH, van Balkom AJLM. Effectiveness of the implementation of guidelines for anxiety disorders in specialized mental health care. *Acta Psychiatr Scand*. 2015;132(1):69-80. doi:10.1111/acps.12338
44. Rush JA, Jain S. Clinical Implications of the STAR*D Trial. In: *Handbook of Experimental Pharmacology*. Springer Nature; 2018:1-49. doi:10.1007/164_2018_153
45. Gustavson K, Knudsen AK, Nesvåg R, Knudsen GP, Vollset SE, Reichborn-Kjennerud T. Prevalence and stability of mental disorders among young adults: Findings from a longitudinal study. *BMC Psychiatry*. 2018;18(1):1-15. doi:10.1186/s12888-018-1647-5
46. Fava M. Diagnosis and Definition of Treatment-Resistant Depression. *Biol Psychiatry*. 2003;53:649-659. doi:10.1016/S0006-3223(03)00231-2
47. Thase ME, Rush AJ. When at First You Don't Succeed: Sequential Strategies for Antidepressant Nonresponders. *J Clin Psychiatry*. 1997;58(suppl 13):23-29.
48. Coles ME, Heimberg RG. Memory biases in the anxiety disorders: Current status. *Clin Psychol Rev*. 2002;22(4):587-627. doi:10.1016/S0272-7358(01)00113-1
49. Sars D, van Minnen A. On the use of exposure therapy in the treatment of anxiety disorders: a survey among cognitive behavioural therapists in the Netherlands. *BMC Psychol*. 2015;3:26. doi:10.1186/s40359-015-0083-2

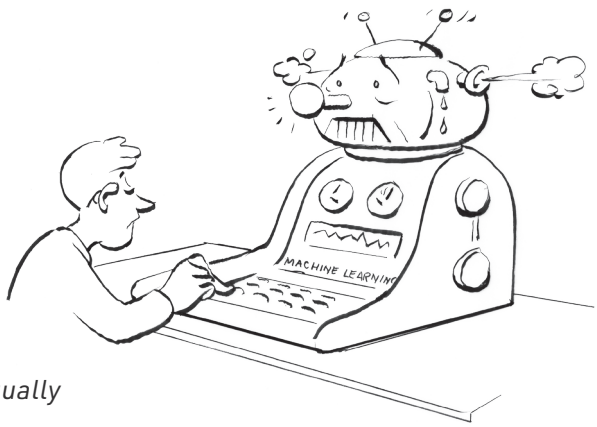


Chapter 7.

Predicting the naturalistic course in anxiety disorders using clinical and biological markers: a machine learning approach

Psychological Medicine (2020)

Wicher A. Bokma §
Paul Zhutovsky §
Erik J. Giltay
Robert A. Schoevers
Brenda W.J.H. Penninx
Anton L.J.M. van Balkom
Neeltje M. Batelaan §
Guido A. van Wingen §



§ *both authors contributed equally*

Abstract

Background

Disease trajectories of patients with anxiety disorders are highly diverse and approximately 60% remain chronically ill. The ability to predict disease course in individual patients would enable personalized management of these patients. This study aimed to predict recovery from anxiety disorders within 2 years by applying a machine learning approach.

Methods

In total, 887 patients with anxiety disorders (panic disorder, generalized anxiety disorder, agoraphobia, or social phobia) were selected from a naturalistic cohort study. A wide array of baseline predictors (N = 569) from five domains (clinical, psychological, socio-demographic, biological, lifestyle) were used to predict recovery from anxiety disorders and recovery from all common mental disorders (CMDs: anxiety disorders, major depressive disorder, dysthymia, or alcohol dependency) at 2-year follow-up using random forest classifiers (RFCs).

Results

At follow-up, 484 patients (54.6%) had recovered from anxiety disorders. RFCs achieved a cross-validated area-under-the-receiving-operator-characteristic-curve (AUC) of 0.67 when using the combination of all predictor domains (sensitivity: 62.0%, specificity: 62.8%) for predicting recovery from anxiety disorders. Classification of recovery from CMDs yielded an AUC of 0.70 (sensitivity: 64.6%, specificity: 62.3%) when using all domains. In both cases, the clinical domain alone provided comparable performances. Feature analysis showed that prediction of recovery from anxiety disorders was primarily driven by anxiety features, whereas recovery from CMDs was primarily driven by depression features.

Conclusions

The current study showed moderate performance in predicting recovery from anxiety disorders over a 2-year follow-up for individual patients and indicates that anxiety features are most indicative for anxiety improvement and depression features for improvement in general.

Introduction

Anxiety disorders are characterized by highly heterogeneous clinical course trajectories. After 2 years, the prognosis varies across disorders with remittance rates of 72.5% for panic disorder without agoraphobia, 69.7% for generalized anxiety disorder, 53.5% for social phobia and 52.7% for panic disorder with agoraphobia (Hendriks, Spijker, Licht, Beekman, & Penninx, 2013). Remitted patients experience a relatively benign course with moderate remaining symptom severity, disability and a low subjective need for care (Batelaan, Rhebergen, Spinhoven, van Balkom, & Penninx, 2014; Spinhoven et al., 2016; van Beljouw, Verhaak, Cuijpers, van Marwijk, & Penninx, 2010). However, around 60% of patients have persistent symptoms, relapses, or chronic disease up to 6 years after the diagnosis (Batelaan et al., 2014; Spinhoven et al., 2016). Disease course in these patients is often characterized by substantial levels of disability. Predicting long-term disease course can be seen as an important step towards personalized medicine (Steyerberg, 2009). This would make targeted treatment efforts viable, in which treatments are tailored towards the individual risk for a poor disease outcome (McGorry, Ratheesh, & O'Donoghue, 2018). However, in anxiety disorders, there is a lack of robust course predictors. For instance, different DSM anxiety disorder diagnoses were shown to be poorly predictive of subsequent course (Batelaan et al., 2014). In current clinical practice, in the absence of valid risk prediction models, course prediction relies solely on clinician's opinions, which show poor accuracy (Randall, Sareen, Chateau, & Bolton, 2019).

Several clinical, psychological, biological, sociodemographic and lifestyle markers are related to the disease course. For instance, higher baseline severity of anxiety symptoms, presence of somatic or psychiatric comorbidity, and higher levels of disability are linked to worse outcomes at 1-year (van Beljouw et al., 2010), 2-year (Batelaan et al., 2014; Hendriks et al., 2013; Scholten et al., 2013), 6-year (Spinhoven et al., 2016), and 12-year follow-up (Bruce et al., 2005). Contrastingly, some authors suggest the same factors lead to better initial treatment results (Baldwin & Tiwari, 2009; Rodriguez et al., 2006). Also, a chronic duration of anxiety was linked to worse outcomes in most studies (Batelaan et al., 2014; Hendriks et al., 2013; Scholten et al., 2013; Spinhoven et al., 2016), while not showing any effect on disease course in another study (Nay, Brown, & Roberson-Nay, 2013). Most studies showed that a younger age at onset was associated with a chronic course (Batelaan et al., 2014; Beesdo-Baum et al., 2012; Rodriguez

et al., 2006), while others showed no such age effect (Nay et al., 2013; Scholten et al., 2013). Inconsistent findings are likely due to methodological differences between studies. Other factors possibly related to worse disease course were duration of untreated illness (Baldwin & Tiwari, 2009), the use of anti-anxiety medication (Bruce et al., 2005; Scholten et al., 2013), and presence of childhood trauma (Asselmann & Beesdo-Baum, 2015; Batelaan et al., 2014; Scholten et al., 2013). Psychological factors that negatively impact anxiety disorder disease course up till 6-year follow-up included high neuroticism (Asselmann & Beesdo-Baum, 2015; Scholten et al., 2013; Spinhoven et al., 2016), low extraversion (Spinhoven et al., 2016), high anxiety sensitivity (Asselmann & Beesdo-Baum, 2015; Scholten et al., 2013), high levels of worrying (Spinhoven et al., 2016), and low mastery (Asselmann & Beesdo-Baum, 2015; Scholten et al., 2013). Only a few studies linked biological parameters to disease course in anxiety disorders: C-reactive protein (CRP) levels were longitudinally associated with anxiety symptoms (Copeland, Shanahan, Worthman, Angold, & Costello, 2012), increasing cortisol levels were linked to higher 6-month anxiety severity in girls (Schiefelbein & Susman, 2006), and lower Brain-Derived Neurotropic Factor (BDNF) levels were found in patients with a poor response to treatment (Kobayashi et al., 2005). However, most research into biological parameters for anxiety disorders was done cross-sectionally, showing that anxiety disorder status is linked to higher CRP-levels (Copeland et al., 2012; Pitsavos et al., 2006; Vogelzangs, Beekman, De Jonge, & Penninx, 2013), higher metabolic syndrome markers (Carroll et al., 2009; Kahl et al., 2015; Perez-Cornago, Ramírez, Zulet, & Martinez, 2014), higher tumour necrosis factor-α (TNF-α) levels (Hoge et al., 2009; Pitsavos et al., 2006), and lower BDNF levels (Molendijk et al., 2012). Inconsistently, anxiety symptoms were linked to both higher (Zoccola, Dickerson, & Yim, 2011) and lower (O 'Donovan et al., 2010) cortisol, as well as higher (Hoge *et al*. 2009; O 'Donovan *et al*. 2010; Pitsavos *et al*. 2006) and lower (Vogelzangs *et al*. 2013) interleukin-6 (IL-6) measurements. Finally, sociodemographic and lifestyle factors such as education years (van Beljouw et al., 2010), age (Asselmann & Beesdo-Baum, 2015; Catarino et al., 2018), partner status (Asselmann & Beesdo-Baum, 2015; Batelaan et al., 2014), social support (van Beljouw et al., 2010), smoking status (Bruce et al., 2005), nicotine dependency (Nay et al., 2013), current financial problems (Nay et al., 2013), employment status (van Beljouw et al., 2010), and income (van Beljouw et al., 2010) were associated with anxiety disorder disease course. In spite of these many variables that predict disease course at the group level, it is not known whether this translates to accurate predictions for individual

patients. Currently, no encompassing model exists with sufficient sensitivity and specificity in disease course prediction to be feasible for use at the level of the individual patient.

A possible explanation for the lack of accuracy in course prediction in anxiety disorders is the complex, multicausal aetiology of anxiety disorders. Univariable and multivariable analyses of predictors of disease course showed low levels of explained variance (Bokma, Batelaan, Hoogendoorn, Penninx, & van Balkom, 2020). Furthermore, the inference is typically done on the group-level which does not allow for generalizable statements for the single individual. Multivariable machine learning (ML) methods provide a possible solution for this problem, as they are well-suited for solving problems with high numbers of predictors in complex, multicausal disorders (Iniesta, Stahl, & McGuffin, 2016). The use of ML in the field of psychiatry may have great potential for its application in the prediction of disease course trajectories (Hahn, Nierenberg, & Whitfield-Gabrieli, 2017). Prediction of the disease course can be regarded as a 'classification' problem, which can be solved using supervised algorithms (Deo, 2015). In these, algorithms are trained on patients with known predictor and outcome variables to derive a function that can be applied to unseen patients to predict their outcome based on the values of their predictor variables. In anxiety disorders, supervised algorithms were applied a few times cross-sectionally, to relate predictors from various domains to current disease status (Woo, Chang, Lindquist, & Wager, 2017) or to predict short-term treatment effects (Lueken & Hahn, 2016). To our best knowledge, however, no studies applied supervised ML algorithms to predict the disease course in anxiety disorders.

The aim of this study was to predict long-term anxiety disorder course, using an ML approach applied to clinical, psychological, biological, sociodemographic and lifestyle baseline data. Specifically, we investigated the utility of a random forest classifier (RFC) (Breiman, 2001) to predict clinical course in patients with any baseline anxiety disorder. Our main outcome was recovery from anxiety disorders at 2-year follow-up. As secondary outcome recovery from all common mental disorders (CMDs) at 2-year follow-up was used. CMDs include anxiety disorders, but also depressive disorders and substance use disorders as these disorders often co-occur, show diagnostic instability over time (Hovenkamp-Hermelink et al., 2016; Lamers et al., 2011; Scholten et al., 2016; Verduijn et al., 2017), and recovery from one but not the other does not index a major improvement

in health. Finally, we assessed which predictor domains contributed most to disease course predictions. We hypothesized that RFCs using a wide array of baseline data from different domains would yield adequate 2-year recovery predictions for both outcomes. Furthermore, we hypothesized that the combination of the five domains would yield the best predictions.

Methods

Study sample

The participants in this study were selected from the multi-site Netherlands Study of Depression and Anxiety (NESDA), an ongoing naturalistic cohort study into the course of depression and anxiety. The baseline sample consists of 2981 participants who were recruited from the community, primary care and specialized mental health care centres. All participants had a lifetime or current depressive disorder or anxiety disorder diagnosis (n = 2329, 78.1%) or were healthy controls (n = 652, 21.9%). NESDA allowed for the presence of comorbid psychiatric disorders, with the exception of psychotic disorders, obsessive compulsive disorder, post-traumatic stress disorder, bipolar disorders, or severe substance use disorders. Exclusion criterion consisted of insufficient proficiency of the Dutch language. Baseline data collection was performed in 2004–2007 and was followed by 1-year, 2-year, 4-year, 6-year, and 9-year follow-up measurements. Full descriptions of the design of NESDA were published previously (Penninx et al., 2008). The study protocol was approved by the Ethical Review Board of all participating institutes and written informed consent was obtained from all participants.

For the purpose of this study, patients with current (6-month) panic disorder (PD, with or without agoraphobia), generalized anxiety disorder (GAD) or social anxiety disorder (SAD) diagnoses at baseline were selected (n = 1206). In our sample, psychiatric comorbidity was allowed. The diagnosis was established according to DSM-IV criteria with the Composite International Diagnostic Interview (CIDI, version 2.1) (American Psychiatric Association, 2000; Wittchen, 1994; World Health Organization, 1998). From these patients, 212 were excluded due to missing diagnostic information at 2-years follow-up. A further 107 patients were removed due to having more than 20% missing variables across predictor variables at baseline. This yielded a final sample of 887 anxiety disorder patients with sufficient data available. Excluded patients showed comparable symptom

severity at baseline – mean anxiety severity (Beck's Anxiety Inventory; BAI): 20.35 ± 11.74 *v.* 18.30 ± 10.48, $t = 1.81$, $p = 0.07$; mean depression severity (Inventory of Depressive Symptomatology-Self Report; IDS-SR): 30.71 ± 12.65 *v.* 29.39 ± 12.65, $t = 0.97$, $p = 0.33$. Excluded patients were younger (mean age: 38.25 ± 12.05 *v.* 41.92 ± 12.20 years, $t = 4.62$, $p < 0.001$), and had a lower mean number of education years: 11.03 ± 3.15 *v.* 11.88 ± 3.35, $t = 3.97$, $p < 0.001$, consistent with differences across the whole NESDA sample (Lamers et al., 2012). Gender did not differ between excluded and included patients (% female in excluded sample 68.2%, in included sample 66.8%, $\chi^2 = 0.22$, $p = 0.64$).

Investigated classifications

Two distinct classification tasks predicting outcomes at 2-year follow-up were performed. Both were binary classification tasks predicting (1) recovery from anxiety disorders or (2) recovery from all CMDs. Anxiety disorders were defined as either PD, agoraphobia, GAD, or SAD. Recovery from anxiety disorders was deemed present if no anxiety disorder diagnoses persisted at follow-up. These diagnoses referred to all follow-up anxiety disorders, not only the index disorder(s). Anxiety disorders, dysthymia, major depressive disorder (MDD) and alcohol dependency are sometimes collectively referred to as CMDs (Ormel et al., 2013; Vollebergh et al., 2001). For the purpose of this study, we defined recovery from all CMDs if at follow-up no anxiety disorders, MDD, dysthymia or alcohol dependency diagnoses were present. Assessment of CMDs is relevant as it is evident from population-based studies that depressive disorders and alcohol dependency are the most commonly occurring comorbidities in anxiety disorders (Alonso & Lépine, 2007; Judd et al., 1998; Wittchen, Kessler, Pfister, & Lieb, 2000), rates of diagnostic instability across anxiety disorders, depressive disorders and alcohol dependency are high (Gustavson et al., 2018; Hovenkamp-Hermelink et al., 2016; Scholten et al., 2016) and recovery from one but not the other does not imply a major improvement in health. We assessed recovery from anxiety disorders as a primary outcome measure and recovery from all CMDs as a secondary outcome measure. These two outcome measures describe recovery in a narrow and a broad perspective (Verduijn et al., 2017).

Baseline predictor variables

At baseline, a wide array of putative predictors from five domains (clinical, psychological, sociodemographic, biological and lifestyle) were selected, yielding a total of 651 variables. In our analyses, only information at the individual item level was used. Total summary scores for questionnaires were not calculated, as these would be correlated to the individual items. The exception was the NEO Five-Factor Inventory (NEO-FFI), as its domains (e.g. neuroticism) are of specific clinical relevance. Items were excluded if more than 20% of patients were missing the corresponding item. This resulted in the inclusion of 569 predictors at baseline (see Table 1). If a variable did not apply for a patient, it was re-coded as a new category for ordinal or nominal variables or as 0 for continuous variables (all continuous variables were positive). Such an encoding allowed to maintain the variable for classification and encoded it with a not naturally occurring value implying that this variable did not apply for this patient. All additional missing variables were imputed using median/mode imputation calculated on the training set (see below) to obtain a full data set. No variable had more than 10% missing values before imputation was applied. Additional information about measurement instruments, variable scoring and collection can be found in the Supplementary Methods. We investigated the predictive capability of all domains individually and the combination of all five domains.

Machine learning algorithm

RFCs (Breiman, 2001) were used in all analyses. RFCs have been shown to perform well on many different machine learning problems (Fernández-Delgado, Cernadas, Barro, & Amorim, 2014), specifically in biomedical sciences (Olson, Cava, Mustahsan, Varik, & Moore, 2018). An RFC is built as an ensemble of many decision trees (Breiman, Friedman, Olshen, & Stone, 1984) which themselves are trained by considering random subsamples of variables and patients for each tree. Such a procedure leads to improved and robust prediction performance in comparison to individual trees (Breiman, 2001). Details on hyperparameters used in the analysis can be found in the Supplementary Methods. All analyses were implemented using the scikit-learn (version 0.20.2) (Pedregosa et al., 2011) and imbalanced-learn toolboxes (version 0.4.3) (Lemaître, Nogueira, & Aridas, 2017) in the Python programming language (version 3.7.2).

Evaluation

To evaluate the performance of our classifiers 10-times-repeated-10-fold-cross-validation was applied. In this procedure, the data set is repeatedly (*n* = 100) divided into disjoint training (90% of data) and test (10% of data) sets and the RFC is only fit on the training data and evaluated on the independent test data. The final performance is obtained as an average across all test set evaluations. We measured performance as area-under-the-receiver-operator-curve (AUC). In addition, we calculated sensitivity, specificity, balanced accuracy – average between sensitivity and specificity – and positive/negative predictive values. To further validate our classification performance label-permutation tests (*n* = 1000) of average AUC values were performed (Ojala & Garriga, 2010). The obtained *p* values were Bonferroni-corrected across five individual and one combination of all domains and alpha was set to 0.05.

To systematically compare the performance of different predictor domains patients were distributed in exactly the same way for each of the classifications, i.e. the train and test set of any cross-validation iteration included the same patients for each predictor domain. This allowed the calculation of normalized average differences in AUC scores across cross-validation iterations for each pair of predictor domains (including the combination of all domains). Non-parametric sign-flipping tests (*n* = 10 000) were then employed to derive *p* values which were Bonferroni-corrected for 30 comparisons with alpha set to 0.05.

Table 1. Included baseline predictor variables across the five predictor domains

Domain	Timespan	Constructs (no of items)	Measurement instruments
Clinical domain (311 variables)	Current	Common mental disorder diagnoses (25), pathological worrying (11), phobic concerns (15), disability (35), all psychotropic medication, by classes (13).	WHO-Composite International Diagnostic Interview (CIDI), Penn State Worry Questionnaire (PSWQ), Fear Questionnaire (FQ), WHO-Disability Assessment Schedule II (WHO-DAS), according to Anatomical Therapeutic Chemical (ATC) codes.
	Past week	Depressive symptoms (28), general distress and somatization (32), mood and anxiety symptoms (30), suicidal ideation (5).	Inventory of Depressive Symptomatology-SR (IDS-SR), Four-Dimensional Symptom Questionnaire (4DSQ), Mood and Anxiety Scoring Questionnaire (MASQ), Suicidal Ideation Scale (SSI).
	Past four weeks	Anxiety symptoms (21), sleep quality (6).	Beck Anxiety Inventory (BAI), Insomnia Rating Scale (ISR),
	Past six months	Perceived need for care (14).	Perceived Need for Care Questionnaire.
	Past three years	Previous psychotropic medication, by classes (6).	According to ATC codes.
	Past four years	Anxiety duration, months (1).	Life Chart Interview (LCI).
	Lifetime	Anxiety and depressive disorders diagnoses (14), bipolar symptoms (13), number of negative life-events (1), childhood trauma (3), convictions about the importance of care and past experiences with care (36).	CIDI, Mood Disorder Questionnaire (MDQ), Brugha questionnaire, NEMESIS questionnaire, QUality Of care Through the Eyes of the patient (QUOTE): Anxiety/Depression version.
Psychological domain (131 variables)	Current	Anxiety sensitivity (16), cognitive reactivity to sadness (34), mastery (5), personality structure according to the Five Factor (76).	Anxiety Sensitivity Index, Leiden Index of Depression Sensitivity, Pearlin Mastery, NEO Five-Factor Inventory.
Sociodemographic domain (71 variables)	Current	Demographic characteristics (6), employment status (5), marital status (2), sexual preference (1), housing status (5), family and household decomposition (6), income (11), religion (1), leisure activities (20), loneliness (11), social support (3).	Self-report questionnaires, de Jong-Gierveld loneliness scale, Close Person Inventory.
Biological domain (49 variables)	Current	Number of chronic diseases (2), chronic pain (1), menstrual cycle status (4), Body Mass Index (1), hip/waist circumference ratio (2), blood pressure (7), handedness (1), hand-grip strength (2), current fever or cold (2), autonomic nervous system function (6), blood plasma measures, including CRP, TNF-α, BDNF, and IL-6 (21).	Chronic graded pain scale, OMRON M4 IntelliSense digital blood pressure monitor, Jamar dynamometer, Vrije Universiteit Ambulatory Measuring System.
Lifestyle domain (7 variables)	Current	Smoking status (1), psychoactive substances use (1), amount of alcohol consumption (1), levels of physical exercise (4).	Fagerström Test for Nicotine Dependence, Alcohol Use Disorders Identification Test, International physical activity questionnaire.

Variable importance

In addition to its strong classification performance RFCs allow to quantify the importance of each variable towards the classification task (Breiman, 2001). However, the standard calculation of variable importance has been shown to be biased (Strobl, Boulesteix, Zeileis, & Hothorn, 2007) and a permutation-based variable importance scheme has been suggested instead (Altmann, Tološi, Sander, & Lengauer, 2010; Hapfelmeier & Ulm, 2013; Strobl et al., 2007). Following this approach, we calculated *p* values for each variable by permuting (n = 1000) every variable separately. The computed *p* values were then corrected according to the false discovery rate (FDR) (Benjamini & Hochberg, 2000) and significance was set to 0.05. Given that variable importance was calculated every cross-validation iteration, important variables were defined as variables which were consistently significant under FDR for at least 50% of all cross-validation iterations. This very stringent procedure for identifying important variables was employed to calculate valid variable importance information specific to the classification task. Variable importance were only investigated for the classifications using the data from the combination of all domains. In addition, we investigated differences in the average rankings of important variables between the two classification tasks. A detailed description of this approach can be found in the Supplementary Methods.

Results

At 2-year follow-up, 484 patients (54.6%) recovered from anxiety disorders, and 362 patients (40.8%) did not have any CMD. Baseline clinical, psychological, sociodemographic, biological and lifestyle variables are provided for patients with and without anxiety disorders at follow-up (Table 2) and for patients with and without CMD at follow-up (online Supplementary Table 1). Various clinical and psychological variables showed differences between the two groups. By contrast, biological and lifestyle status did not differ between the two groups.

Recovery from anxiety disorders

Classification performance

Results of our evaluation of the RFC when predicting recovery from anxiety disorders are reported in Table 3 and Fig. 1*A*. AUC values for the predictor domains ranged from 0.49 to 0.67 with significant ($p_{Bonferroni} < 0.05$) AUC values obtained for the clinical (0.67), and psychological (0.65) domains, as well as for the combination of all domains (0.67). Classification accuracies were small to moderate with the highest accuracy achieved by the combination of all domains (62.4%) with a sensitivity of 62.0% and specificity of 62.8%. In addition, we investigated the performance of the RFC for subgroups of patients who had any comorbidity (MDD, dysthymia, or alcohol dependency, n = 252 recovered, n = 248 persistent) at baseline and for patients who did not (n = 232 recovered, n = 155 persistent). For that, the RFC trained on all data domains and all patients of the training set was evaluated within the two subgroups on the test set separately. The RFC obtained an average AUC of 0.64 within the no-comorbidity group and an AUC of 0.68 within the comorbidity group showing slightly increased performance for predictions within the comorbidity group.

Domain comparisons

When comparing different domains according to their AUC a clear ordering was observed: The clinical domain outperformed every other domain except for the combination of all domains ($p_{Bonferroni} < 0.05$), the psychological domain outperformed the sociodemographic, biological, and lifestyle domains ($p_{Bonferroni} < 0.05$), the sociodemographic domain outperformed the biological and lifestyle domains ($p_{Bonferroni} < 0.05$), and the biological domain outperformed the lifestyle domain ($p_{Bonferroni} < 0.05$). The combination of all domains was better than any domain except for the clinical domain ($p_{Bonferroni} < 0.05$).

Variable importance

Consistently selected significant variables ($N = 17$) identified through a permutation-based variable importance calculation of the RFC are reported in online Supplementary Table 2. Only variables from the clinical and psychological domain were selected. These variables were derived from different measurement instruments (BAI, IDS-SR, Fear Questionnaire (FQ), NEO-FFI, WHO-Disability Assessment (WHO-DAS), Four-Dimensional Symptom Questionnaire (4DSQ), Mastery scale) but all referred to characteristic anxiety symptoms, with an emphasis on anxious arousal items.

Recovery from all common mental disorders

Classification performance

Results of the second classification procedure predicting recovery from CMDs are reported in Table 4 and Fig. 1*B*. AUC values ranged from 0.53 to 0.70 with significant ($p_{Bonferroni} < 0.05$) AUC values obtained for the clinical (0.70), psychological (0.67), and sociodemographic domain (0.65) as well as the combination of all domains (0.70). The highest accuracy was achieved by the combination of all domains (63.4%) with a sensitivity of 64.6% and a specificity of 62.3%. As in the case of the prediction of the recovery from anxiety disorders, we investigated the performance of the RFC for subgroups of patients who had (n = 164 recovered, n = 336 persistent) or did not (n = 198 recovered, n = 189 persistent) have any comorbidities at baseline. For that, the RFC trained on the combintation of all domains and all patients of the training set was evaluated within the two sub-groups on the test set separately. The RFC obtained an AUC of 0.62 within the no-comorbidity group and an AUC of 0.73 within the comorbidity group. As in the case of the prediction of recovery from anxiety disorders the RFC was showing better performance for patients with comorbidities at baseline.

Domain comparisons

The best performing domains for this classification were the same as in the recovery from anxiety disorders classification. The clinical domain and the combination of all domains did not differ in their performance but outperformed any other domain during the classification. The order for the performance of the other domains was the same as with the recovery from anxiety disorders classification.

Variable importance

48 variables were identified as being consistently selected significant variables contributing to the classification (online Supplementary Table 3). In this classification, selected variables included a larger set of measures related to mood disorders and not only anxiety symptomatology. With one exception (sociodemographic) all variables were again selected from the clinical or psychological domain.

Difference in important variables between prediction analyses

Variables which were more (or less) important in the prediction of recovery from anxiety disorders than the prediction of all CMDs are reported in online Supplementary Table 4. These results confirmed the importance of

anxiety-related variables for the prediction of recovery from anxiety, and the importance of depression-related variables for the prediction of recovery from all CMDs.

Table 2. Baseline characteristics of anxiety disorder sample, group comparisons between patients who had no anxiety disorder (*n* = 484) at 2-year follow-up and patients who did (*n* = 403)

Two-year anxiety disorder status	Recovered (n=484)	Persistent (n=403)	Statistics	p
Clinical domain				
PD diagnosis	176 (36.4%)	192 (47.6%)	$X^2 = 11.52$	**<0.001**
Agoraphobia diagnosis	141 (29.1%)	176 (43.7%)	$X^2 = 20.24$	**<0.001**
SAD diagnosis	196 (40.5%)	212 (53.6%)	$X^2 = 12.98$	**<0.001**
GAD diagnosis	141 (29.1%)	136 (33.7%)	$X^2 = 2.18$	0.14
MDD diagnosis	174 (36.0%)	188 (46.7%)	$X^2 = 10.42$	**0.001**
Dysthymia diagnosis	58 (12.0%)	88 (21.8%)	$X^2 = 15.53$	**<0.001**
Use of psychotropic medication, *current*	345 (71.3%)	294 (73.0%)	$X^2 = 0.31$	0.58
Avoidance behaviour severity, *mean FQ*, *current*	31.76 ± 18.21	40.90 ± 20.07	t = -6.98	**<0.001**
Pathological worrying severity, *mean PSWQ*, *current*	35.95 ± 9.91	39.56 ± 9.39	t = -5.52	**<0.001**
Suicidal thoughts, *SSI*, *past week*	72 (14.9%)	111 (27.5%)	$X^2 = 21.55$	**<0.001**
Level of distress, *mean 4DSQ*, *past week*	16.03 ± 8.94	19.85 ± 8.75	t = -6.39	**<0.001**
Depressive symptoms severity, *mean IDS-SR*, *past week*	26.58 ± 12.00	32.78 ± 12.59	t = -7.48	**<0.001**
Sleep disturbances, *mean ISR, past four weeks*	9.39 ± 5.15	10.19 ± 5.24	t = -2.28	**0.02**
Anxiety symptoms severity, *mean BAI*, *past month*	15.97 ± 9.36	21.10 ± 11.05	t = -7.49	**<0.001**
Percentage of time spent with anxiety symptoms, *LCI, past four years*	43.81% ± 33.20	54.04% ± 34.20	t = -4.37	**<0.001**
History of childhood life events[1]	89 (18.4%)	74 (18.4%)	$X^2 = 0.00$	0.99
History of childhood trauma[2]	258 (53.3%)	247 (61.4%)	$X^2 = 5.93$	**0.02**
History of serious suicide attempts	66 (13.7%)	87 (21.6%)	$X^2 = 9.57$	**0.002**
Psychological domain				
Neuroticism, *mean NEO-FF subscale*	40.46 ± 6.89	43.66 ± 6.73	t = -6,95	**<0.001**
Extraversion, *mean NEO-FFI subscale*	34.70 ± 6.51	32.43 ± 6.82	t = 5.06	**<0.001**
Conscientiousness, *mean NEO-FFI subscale*	40.88 ± 6.45	39.23 ± 6.36	t = 3.82	**<0.001**
Agreeableness, *mean NEO-FFI subscale*	43.38 ± 5.37	42.59 ± 5.27	t = 2.20	**0.03**
Openness, *mean NEO-FFI subscale*	38.25 ± 6.03	38.04 ± 6.32	t = 0.51	0.61
Cognitive reactivity to sadness, *mean LEIDS*	40.80 ± 17.98	46.76 ± 17.98	t = -4.91	**<0.001**
Anxiety sensitivity, *mean ASI*	33.63 ± 9.47	36.58 ± 10.43	t = -4.35	**<0.001**
Mastery, *mean Mastery scale*	15.77 ± 4.02	13.89 ± 4.06	t = 6.90	**<0.001**
Sociodemographic domain				
Age in years	41.88 ± 12.09	41.97 ± 12.34	t = -0.11	0.91
Education years	12.02 ± 3.29	11.70 ± 3.41	t = 1.43	0.15
Female gender	329 (68.0%)	276 (68.5%)	$X^2 = 0.03$	0.87
Currently employed	280 (57.9%)	206 (51.1%)	$X^2 = 4.03$	**0.05**

Table continues

Table 2. Continued

	Has children	268 (55.4%)	212 (52.6%)	$X^2 = 0.68$	0.41
	Current severe loneliness	47 (9.7%)	58 (14.4%)	$X^2 = 4.57$	**0.03**
Biological domain					
	Number of chronic somatic diseases	0.67 ± 0.89	0.72 ± 0.95	t = -0.82	0.41
	Chronic pain with high disability	100 (20.7%)	121 (30.0%)	$X^2 = 10.31$	**0.001**
	BMI	25.46 ± 4.72	25.71 ± 5.52	t = -0.74	0.46
	Mean heart rate (bpm)	71.70 ± 9.59	72.05 ± 10.12	t = -0.52	0.60
	Systolic blood pressure (mmHg)	136.3 ± 20.63	135.9 ± 17.97	t = 0.31	0.76
	CRP (mg/L, n=876)	2.67 ± 4.05	3.12 ± 6.29	t = -1.26	0.21
	IL-6 (pg/ml, n=876)	1.28 ± 3.00	1.43 ± 3.15	t = -0.69	0.49
	TNF-α (pg/ml, n=871)	1.07 ± 1.28	1.04 ± 1.12	t = 0.41	0.69
	BDNF (ng/ml, n=865)	9.18 ± 3.64	9.20 ± 3.46	t = -0.08	0.94
Lifestyle domain					
	Former smoker	153 (31.6%)	119 (29.5%)	$X^2 = 2.86$	0.24
	Current smoker	174 (36.0%)	167 (41.4%)		
	Low physical activity, *past week*	103 (22.7%)	98 (25.3%)	$X^2 = 1.84$	0.40
	High physical activity, *past week*	156 (34.4%)	117 (30.2%)		
	Any substance use, *past week*	33 (6.8%)	33 (8.2%)	$X^2 = 0.60$	0.44
	Hazardous drinking or alcohol dependency,[3] *past year*	109 (22.6%)	87 (21.6%)	$X^2 = 0.13$	0.71

PD, panic disorder; SAD, social anxiety disorder; GAD, generalized anxiety disorder; MDD, major depressive disorder; FQ, Fear Questionnaire; PSWQ, Penn State Worry Questionnaire; SSI, Suicidal Ideation Scale; 4DSQ, Four-Dimensional Symptom Questionnaire; IDS-SR, Inventory of Depressive Symptomatology-SR; ISR, Insomnia Rating Scale; BAI, Beck's Anxiety Inventory; LCI, life chart interview; NEO-FFI, NEO Five-Factor Inventory; LEIDS, Leiden Index of Depression Sensitivity; ASI, Anxiety Sensitivity Index; BMI, Body Mass Index; CRP, c-reactive protein; IL-6, interleukin-6; TNF-α, tumour necrosis factor-α; BDNF, Brain-Derived Neurotrophic Factor.

p values shown in bold are <0.05.

[a]Childhood life events (<16 years of age) were parental divorce, being placed in a juvenile prison, raised in a foster family, placed in a child home, death of a parent.

[b]Childhood trauma included emotional neglect, psychological abuse, physical abuse, and sexual abuse.

[c]As measured with the AUDIT. Scores above 8 are reflective of hazardous drinking, scores at 13 or higher (females) and 15 or higher (males) are indicative of probable alcohol dependency.

Transfer analysis

We replicated the classification of recovery from anxiety disorders at 2-year follow-up in a transfer learning setting: in such an approach we utilized the labels indicating recovery of CMDs during the training of the RFC classifier (training set) but subsequently evaluated its performance on the test set using the recovery from anxiety disorder labels. The result of this analysis can be seen in online Supplementary Table 5. Utilizing the transfer learning approach led to improved performance in predicting anxiety disorder recovery (AUC = 0.71 *v.* AUC = 0.67 for both training and testing on anxiety disorder recovery labels using either only the clinical or the combination of all domains). The increased performance was observed due to an increase

in sensitivity of the classification for correctly identifying recovered anxiety patients. For all individual domains and the combination of them, sensitivity increased by 7.6 ± 1.9 when training on the CMDs labels first. Specificity only decreased slightly (mean decrease: 2.7 ± 0.8) which led to the improved overall performance.

Discussion

One of the most important goals in personalized medicine is providing individual disease course predictions. Our results show that individual prediction of 2-year course in anxiety disorders is possible using various predictors but it is only moderately successful. The main outcome measure was recovery from anxiety disorders and our predictions reached a balanced accuracy of 62.4% with an AUC of 0.67. The current performance by itself does not warrant implementation of our models in routine psychiatric care as it would yield too many false positives/negatives. However, predictive properties of clinician opinion in predicting disease course in anxiety disorders are not available and therefore it remains unclear which predictive performance threshold is needed for a statistical model to surpass clinician opinion and become an improvement over current routine care.

Our study yielded two models with comparable accuracy for predicting 2-year anxiety disorder course: one consisting of predictors from all five domains and one consisting of predictors only from the clinical domain. Biological, lifestyle, and sociodemographic predictors did not contribute significantly to course prediction. This is surprising as these domains were previously shown to be related to anxiety disorder aetiology. Our results thereby suggest that the underlying aetiology is of less importance to course prediction after the development of threshold disorders and that after anxiety disorders have developed, phenotypical characteristics have more impact on subsequent disease course. This is evident from the individual features that contributed most to the classification. All of these features reflected symptoms, psychological states or traits associated with the emotions of fear and anxiety, such as the presence of 'phobic symptoms', difficulty 'walking alone in a busy street' or 'dealing with people you don't know', 'feeling tense', 'not liking to be where the action is', and 'feeling faint or lightheaded'. A previous NESDA study that aimed to predict the naturalistic course in depression showed similar performance to the current study when 2-year follow-up MDD diagnosis was correctly classified with an AUC of

0.66 and balanced accuracy of 62% (Dinga et al., 2018). In this study, clinical features were most important as well, though the nature of those items was related to depression.

Table 3. Evaluation of the 2-year recovery from anxiety disorders classification [mean (S.D.)]

Domains	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV
Clinical	**0.67 (0.05)***	61.7 (4.4)	61.5 (6.3)	61.9 (7.6)	0.66 (0.05)	0.57 (0.04)
Psychological	**0.65 (0.05)***	61.0 (4.5)	60.0 (6.4)	61.9 (7.5)	0.66 (0.05)	0.56 (0.05)
Socio-demographic	0.56 (0.06)	53.1 (5.1)	49.7 (7.4)	56.5 (7.5)	0.58 (0.06)	0.48 (0.05)
Biological	0.53 (0.06)	52.7 (4.9)	50.3 (6.8)	55.0 (7.6)	0.57 (0.05)	0.48 (0.05)
Lifestyle	0.49 (0.05)	50.2 (4.3)	46.6 (5.5)	53.7 (7.6)	0.55 (0.05)	0.46 (0.04)
Combination	**0.67 (0.05)***	62.4 (4.6)	62.0 (6.1)	62.8 (7.5)	0.67 (0.05)	0.58 (0.05)

AUC, area-under-receiver-operator-curve; PPV, positive predictive value; NPV, negative predictive value; * $p_{Bonferroni} < 0.05$.
p values shown in bold are <0.05.

Table 4. Evaluation of the two-year recovery from all common mental disorders classification [mean (SD)]

Domains	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV
Clinical	**0.70 (0.05)***	62.2 (4.6)	65.0 (7.1)	59.3 (5.6)	0.52 (0.05)	0.71 (0.05)
Psychological	**0.67 (0.05)***	62.2 (4.8)	61.8 (8.4)	62.6 (6.4)	0.53 (0.05)	0.71 (0.05)
Socio-demographic	**0.65 (0.05)***	60.8 (5.2)	65.2 (7.5)	56.5 (6.5)	0.51 (0.05)	0.70 (0.05)
Biological	0.57 (0.05)	56.0 (4.8)	57.5 (8.2)	54.6 (6.7)	0.47 (0.05)	0.65 (0.05)
Lifestyle	0.53 (0.05)	51.8 (4.7)	62.3 (7.9)	41.2 (6.5)	0.42 (0.04)	0.61 (0.06)
Combination	**0.70 (0.05)***	63.4 (4.8)	64.6 (7.3)	62.3 (6.1)	0.54 (0.05)	0.72 (0.05)

AUC, area-under-receiver-operator-curve; PPV, positive predictive value; NPV, negative predictive value; * $p_{Bonferroni} < 0.05$.
p values shown in bold are <0.05.

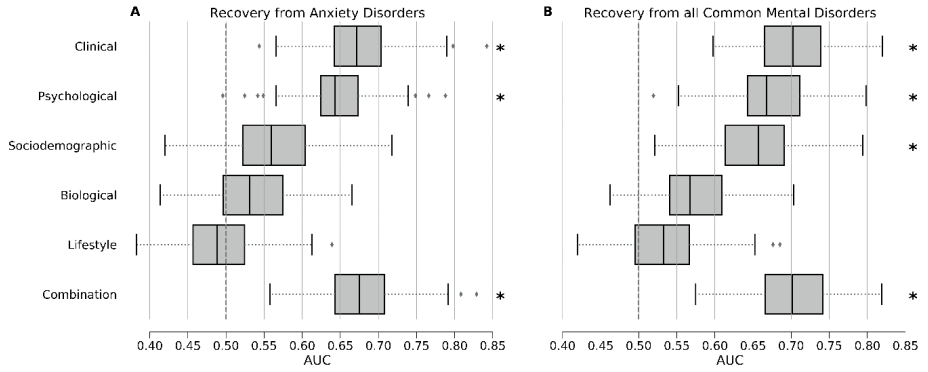


Figure 1. Classification performance of random forest classifiers. Performance is quantified by area-under-the-receiver-operator-curve (AUC) values calculated for each test set of all cross-validation iterations and is shown in box-and-whisker plots for all data domains. (a) Performance of the recovery from anxiety disorders prediction, (b) Performance of the recovery from all common mental disorders prediction. Asterisks mark a significant classification performance according to label-permutation tests (n = 1000) and Bonferroni-correction for six tests.The dashed line indicates chance-level performance.

As anxiety disorders and other psychiatric disorders frequently co-occur and show diagnostic instability over time, a secondary outcome was assessed. This broad perspective model was trained on recovery from all CMDs and showed marginally higher accuracy (63.4%) and AUC (0.70) in comparison with the main narrow perspective outcome. Like in the narrow perspective, omitting all domains except the clinical domain did not lead to a significant loss of predictive power (accuracy = 62.2% and AUC = 0.70).

The individual features that were most consistently chosen during the classification again were almost exclusively from the clinical and psychological domains. Symptoms, psychological traits, and psychological states associated with depression and worrying contributed most to the classification. For instance: 'feeling down', 'feeling sad', having 'a desire to die', 'suffering from worry', 'feeling tense', and 'having little control about the things that happen'. This suggests that predictions for recovery from all CMDs were largely driven by co-occurring depressive symptoms. Our decision to investigate the CMDs classification was also supported by the results of the additional transfer analysis which showed improved performance (accuracy = 63.3% and AUC = 0.71 for the combination of all domains data) when using the recovery from all CMDs labelling during training and the recovery from anxiety labels during model evaluation. This

analysis showed that patients suffering from any mental disorder at 2-year follow-up – anxiety or not – constituted a more homogenous group while patients who fully recovered were more easily identified than patients only recovering from anxiety disorders (but having an additional CMD instead). This suggests that applying a broad perspective in future attempts in clinical prediction is more feasible for anxiety disorders.

Previous ML studies in anxiety disorders were invariably small in sample size and most focused on predicting immediate treatment response using neuroimaging data (Ball, Stein, Ramsawh, Campbell-Sills, & Paulus, 2014; Doehrmann et al., 2013; Hahn et al., 2014; Pantazatos, Talati, Schneier, & Hirsch, 2014; Whitfield-Gabrieli et al., 2016). Some studies used clinical, biological and/or neuroimaging data to distinguish between different types of anxiety disorders and healthy controls (Carpenter, Sprechmann, Calderbank, Sapiro, & Egger, 2016; Frick et al., 2014; Hilbert, Lueken, Muehlhan, & Beesdo-Baum, 2017; Pantazatos et al., 2014). To the best of our knowledge, this is the first study into individual long-term course prediction in anxiety disorders. A strength of this study is the use of a large dataset with a high number of variables from a variety of predictor domains, most of which were previously related to disease course at the group level. In addition, using RFCs allowed for combining large numbers of predictors into an overall model and allowed the identification of the most contributing predictors, providing insight into the possible processes involved with recovery in anxiety disorders.

In spite of the wide array of predictors, the current study showed only moderate accuracy. This has a number of explanations. First, NESDA is a naturalistic cohort study in which the exposure to environmental stressors and treatment regimens varied across patients during the 2-year follow-up period. These different exposures will have impacted the 2-year outcomes. Furthermore, different data types might improve predictive accuracy. For instance, previous ML studies showed the strong potential of neuroimaging data to predict treatment response in anxiety disorders (Ball et al., 2014; Doehrmann et al., 2013; Hahn et al., 2014; Pantazatos et al., 2014; Whitfield-Gabrieli et al., 2016), sometimes exceeding predictions made using clinical data (Ball et al., 2014; Doehrmann et al., 2013). Our study did not encompass neuroimaging data, as these were only available in a subset of NESDA participants (Janssen, Mourão-Miranda, & Schnack, 2018). Other examples include gait analysis (Zhao et al., 2019), actigraphy (Merikangas et al., 2019), or social media data (Reece & Danforth, 2017). Additionally,

more frequent data collection might improve predictive accuracy (Kubben, Dumontier, & Dekker, 2019), which has now been implemented in the most recent wave of NESDA (Difrancesco et al., 2019). However, it is worth noting that our analyses showed that using a large set of variables from various domains (either combined or independently) did not outperform the clinical domain alone. Finally, future studies could explore differences in predictive performance across different patient subgroups, by analyzing separate patient groups consisting of different anxiety disorders, or groups with different comorbidity patterns separately.

Clinical care for anxiety disorders would benefit greatly from improved course prediction as it would pave the way for targeted treatments. The current study showed moderate accuracy in predicting recovery from anxiety disorders over a 2-year follow-up for individual patients. Items from the clinical and psychological domain were the most contributing predictors, while biological, lifestyle, and sociodemographic predictors were contributing less. The limited performance while using a wide array of predictors does not justify application in routine clinical care. The results from our study can, however, be used as a benchmark for future studies, with future studies likely resulting in further enhancements of the predictive properties. It has long been argued that statistical modelling will exceed clinician opinion in prediction problems (Ayres, 2007; Meehl, 1954), with clinician interpretation of statistical models likely yielding the best predictive power (Kuhn & Johnson, 2013). As a result, statistical models will increasingly become an addition to clinician opinion. Eventually, targeted treatment regimens and secondary prevention strategies will become more feasible if predictive models further evolve. This study provides an important first step towards valid long-term ML-based predictions in anxiety disorders.

References

Alonso, J., & Lépine, J. (2007). Overview of key data from the European Study of the Epidemiology of Mental Disorders (ESEMeD). *Journal of Clinical Psychiatry*, *68*(suppl 2), 3–9.

Altmann, A., Toloși, L., Sander, O., & Lengauer, T. (2010). Permutation importance: A corrected feature importance measure. *Bioinformatics (Oxford, England)*, *26*(10), 1340–1347. doi:10.1093/bioinformatics/btq134.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders DSM-IV-TR* (4th ed). New York, NY, US: American Psychiatric Association.

Asselmann, E., & Beesdo-Baum, K. (2015). Predictors of the course of anxiety disorders in adolescents and young adults. *Current Psychiatry Reports*, *17* (2), 1–8. doi:10.1007/s11920-014-0543-z.

Ayres, I. (2007). *Super crunchers: Why thinking-By-numbers is the new way tobe smart*. New York, NY, US: Bantam.

Baldwin, D. S., & Tiwari, N. (2009). The pharmacologic treatment of patients with generalized anxiety disorder: Where are we now and where are we going? *CNS Spectrums*, *14*(2 Suppl 3), 5–12.

Ball, T. M., Stein, M. B., Ramsawh, H. J., Campbell-Sills, L., & Paulus, M. P. (2014). Single-subject anxiety treatment outcome prediction using functional neuroimaging. *Neuropsychopharmacology*, *39*(5), 1254–1261. doi:10.1038/npp.2013.328.

Batelaan, N. M., Rhebergen, D., Spinhoven, P., van Balkom, A. J., & Penninx,

B. W. J. H. (2014). Two-Year course trajectories of anxiety disorders: Do DSM classifications matter? *The Journal of Clinical Psychiatry*, *75*(09), 985–993. doi:10.4088/JCP.13m08837.

Beesdo-Baum, K., Knappe, S., Fehm, L., Höfler, M., Lieb, R., Hofmann, S. G.,& Wittchen, H. U. (2012). The natural course of social anxiety disorder among adolescents and young adults. *Acta Psychiatrica Scandinavica*, *126*(6), 411–425. doi:10.1111/j.1600-0447.2012.01886.x.

Benjamini, Y., & Hochberg, Y. (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. *Journal of Educational and Behavioral Statistics*, *25*(1), 60–83. doi:10.3102/ 10769986025001060.

Bokma, W. A., Batelaan, N. M., Hoogendoorn, A. W., Penninx, B. W. J. H., & van Balkom, A. J. L. M. (2020). A clinical staging approach to improving diagnostics in anxiety disorders: Is it the way to go? *Australian & New Zealand Journal of Psychiatry*, *54*(2), 173–184.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32. doi:10.1023/A:1010933404324.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software.

Bruce, S. E., Yonkers, K. A., Otto, M. W., Eisen, J. L., Weisberg, R. B., Pagano, M., ... Keller, M. B. (2005). Influence of psychiatric comorbidity on recovery and recurrence in generalized anxiety disorder, social phobia, and panic disorder: A 12-year prospective study. *American Journal of Psychiatry*, *162*(6), 1179–1187. doi:10.1176/appi.ajp.162.6.1179.

Carpenter, K. L. H., Sprechmann, P., Calderbank, R., Sapiro, G., & Egger, H. L. (2016). Quantifying risk for anxiety disorders in preschool children: A machine learning approach. *PLoS One*, *11*(11), 1–20. doi:10.7910/DVN/N42LWG.

Carroll, D., Phillips, A. C., Thomas, G. N., Gale, C. R., Deary, I., & Batty, G. D. (2009). Generalized anxiety disorder is associated with metabolic syndrome in the Vietnam experience study. *Biological Psychiatry*, *66*(1), 91–93. doi:10.1016/j.biopsych.2009.02.020.

Catarino, A., Bateup, S., Tablan, V., Innes, K., Freer, S., Richards, A., ... Blackwell, A. D. (2018). Demographic and clinical predictors of response to internet-enabled cognitive–behavioural therapy for depression and anxiety. *BJPsych Open*, *4*(5), 411–418. doi:10.1192/bjo.2018.57.

Copeland, W. E., Shanahan, L., Worthman, C., Angold, A., & Costello, E. J. (2012). Generalized anxiety and C-reactive protein levels: A prospective, longitudinal analysis. *Psychological Medicine*, *42*(12), 2641–2650. doi:10.1017/S0033291712000554.

Deo, R. C. (2015). Machine learning in medicine. *Circulation*, *132*(20), 1920– 1930. doi:10.1161/CIRCULATIONAHA.115.001593.

Difrancesco, S., Lamers, F., Riese, H., Merikangas, K. R., Beekman, A. T. F., Hemert, A. M., ... Penninx, B. W. J. H. (2019). Sleep, circadian rhythm, and physical activity patterns in depressive and anxiety disorders: A 2-week ambulatory assessment study. *Depression and Anxiety*, *36*, 975– 986. 10.1002/da.22949.

Dinga, R., Marquand, A. F., Veltman, D. J., Beekman, A. T. F., Schoevers, R. A., van Hemert, A. M., ... Schmaal, L. (2018). Predicting the naturalistic course of depression from a wide range of clinical, psychological, and biological data: A machine learning approach. *Translational Psychiatry*, *8*(1), 241. doi:10.1038/s41398-018-0289-1.

Doehrmann, O., Ghosh, S. S., Polli, F. E., Reynolds, G. O., Horn, F., Keshavan, A., ... Gabrieli, J. D. (2013). Predicting treatment response in social anxiety disorder from functional magnetic resonance imaging. *JAMA Psychiatry*, *70* (1), 87–97. doi:10.1001/2013.jamapsychiatry.5.

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research*, *15*, 3133–3181.

Frick, A., Gingnell, M., Marquand, A. F., Howner, K., Fischer, H., Kristiansson, M., ... Furmark, T. (2014). Classifying social anxiety disorder using multivoxel pattern analyses of brain function and structure. *Behavioural Brain Research*, *259*, 330–335. doi:10.1016/j.bbr.2013.11.003.

Gustavson, K., Knudsen, A. K., Nesvåg, R., Knudsen, G. P., Vollset, S. E., & Reichborn-Kjennerud, T. (2018). Prevalence and stability of mental disorders among young adults: Findings from a longitudinal study. *BMC Psychiatry*, *18*(1), 1–15. doi:10.1186/s12888-018-1647-5.

Hahn, T., Kircher, T., Straube, B., Wittchen, H.-U., Konrad, C., Ströhle, A., ... Lueken, U. (2014). Predicting treatment response to cognitive behavioral therapy in panic disorder with agoraphobia by integrating local neural information. *JAMA Psychiatry*, *72*, 68–74. 10.1001/jamapsychiatry.2014.1741.

Hahn, T., Nierenberg, A. A., & Whitfield-Gabrieli, S. (2017). Predictive analytics in mental health: Applications, guidelines, challenges and perspectives. *Molecular Psychiatry*, *22*(1), 37–43. doi:10.1038/mp.2016.201.

Hapfelmeier, A., & Ulm, K. (2013). A new variable selection approach using Random Forests. *Computational Statistics and Data Analysis*, *60*(1), 50– 69. doi:10.1016/j.csda.2012.09.020.

Hendriks, S. M., Spijker, J., Licht, C. M. M., Beekman, A. T. F., & Penninx, B. W. J. H. (2013). Two-year course of anxiety disorders: Different across disorders or dimensions? *Acta Psychiatrica Scandinavica*, *128*(3), 212– 221. doi:10.1111/acps.12024.

Hilbert, K., Lueken, U., Muehlhan, M., & Beesdo-Baum, K. (2017). Separatinggeneralized anxiety disorder from major depression using clinical, hormonal, and structural MRI data: A multimodal machine learning study. *Brain and Behavior*, *7*(3), 1–11. doi:10.1002/brb3.633.

Hoge, E. A. A., Brandstetter, K., Moshier, S., Pollack, M. H. H., Wong, K. K. K.,& Simon, N. M. M. (2009). Broad spectrum of cytokine abnormalities in panic disorder and posttraumatic stress disorder. *Depression and Anxiety*, *26*, 447–455. doi:10.1002/da.20564.

Hovenkamp-Hermelink, J. H. M., Riese, H., Van Der Veen, D. C., Batelaan, N.M., Penninx, B. W. J. H., & Schoevers, R. A. (2016). Low stability of diagnostic classifications of anxiety disorders over time: A six-year follow-up of the NESDA study. *Journal of Affective Disorders*, *190*, 310–315. doi:10.1016/j.jad.2015.10.035.

Iniesta, R., Stahl, D., & McGuffin, P. (2016). Machine learning, statistical learning and the future of biological research in psychiatry. *Psychological Medicine*, *46*, 2455–2465. 10.1017/S0033291716001367.

Janssen, R. J., Mourão-Miranda, J., & Schnack, H. G. (2018). Making individual prognoses in psychiatry using neuroimaging and machine learning. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(9), 798–808. 10.1016/j.bpsc.2018.04.004.

Judd, L. L., Kessler, R. C., Paulus, M. P., Zeller, P. V., Wittchen, H. U., & Kunovac, J. L. (1998). Comorbidity as a fundamental feature of generalized anxiety disorders: Results from the National Comorbidity Study (NCS). *Acta Psychiatrica Scandinavica. Supplementum*, *393*, 6–11.

Kahl, K. G., Schweiger, U., Correll, C., Müller, C., Busch, M. L., Bauer, M., & Schwarz, P. (2015). Depression, anxiety disorders, and metabolic syndromein a population at risk for type 2 diabetes mellitus. *Brain and Behavior*, *5*(3),e00306. doi:10.1002/brb3.306.

Kobayashi, K., Shimizu, E., Hashimoto, K., Mitsumori, M., Koike, K., Okamura, N., ... Iyo, M. (2005). Serum brain-derived neurotrophic factor (BDNF) levels in patients with panic disorder: As a biological predictorof response to group cognitive behavioral therapy. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *29*(5), 658–663. doi: 10.1016/j.pnpbp.2005.04.010.

Kubben, P., Dumontier, M., & Dekker, A. (2019). *Fundamentals of clinical data science*. Cham, Switzerland: Springer Open.

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (5th ed.). New York: Springer. doi:10.1007/978-1-4614-6849-3.

Lamers, F., Hoogendoorn, A. W., Smit, J. H., van Dyck, R., Zitman, F. G., Nolen, W. A., & Penninx, B. W. (2012). Sociodemographic and psychiatricdeterminants of attrition in the Netherlands Study of Depression and Anxiety (NESDA). *Comprehensive Psychiatry*, *53*(1), 63–70. doi:10.1016/ j.comppsych.2011.01.011.

Lamers, F., van Oppen, P., Comijs, H. C., Smit, J. H., Spinhoven, P., van Balkom, A. J. L. M., ... Penninx, B. W. J. H. (2011). Comorbidity patternsof anxiety and depressive disorders in a large cohort study: The Netherlands Study of Depression and Anxiety (NESDA). *The Journal of Clinical Psychiatry*, *72*(3), 341–348. doi:10.4088/ JCP.10m06176blu.

Lemaître, G., Nogueira, F., & Aridas, C. K. (2017). Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning. *Journal of Machine Learning Research*, *18*, 1–5. Retrieved from http:// www.jmlr.org/papers/ volume18/16-365/16-365.pdf.

Lueken, U., & Hahn, T. (2016). Functional neuroimaging of psychotherapeuticprocesses in anxiety and depression: From mechanisms to predictions. *Current Opinion in Psychiatry*, *29*(1), 25–31. doi:10.1097/ YCO.0000000000000218.

McGorry, P. D., Ratheesh, A., & O'Donoghue, B. (2018). Early intervention– an implementation challenge for 21st century Mental Health Care. *JAMA Psychiatry*, *75*(6), 545–546. doi:10.1001/jamapsychiatry.2018.0621.

Meehl, P. E. (1954). *Clinical. v. statistical prediction*. Minneapolis, MN, US: University of Minnesota.

Merikangas, K. R., Swendsen, J., Hickie, I. B., Cui, L., Shou, H., Merikangas, A. K., ... Zipunnikov, V. (2019). Real-time mobile monitoring of the dynamic associations among motor activity, energy, mood, and sleep in adults with bipolar disorder. *JAMA Psychiatry*, *Feb*, 190–198. doi:10.1001/jamapsychiatry.2018.3546.

Molendijk, M. L., Bus, B. A., Spinhoven, P., Penninx, B. W., Prickaerts, J., OudeVoshaar, R. C., & Elzinga, B. M. (2012). Gender specific associations of serum levels of brain-derived neurotrophic factor in anxiety. *World Journal of Biological Psychiatry*, *13*(7), 535–543. doi:10.3109/15622975.2011.587892.

Nay, W., Brown, R., & Roberson-Nay, R. (2013). Longitudinal course of panic disorder with and without agoraphobia using the national epidemiologic survey on alcohol and related conditions (NESARC). *Psychiatry Research*,*208*(1), 54–61. doi:10.1016/j.psychres.2013.03.006.

O 'Donovan, A., Hughes, B. M., Slavich, G. M., Lynch, L., Cronin, M.-T., O'Farrelly, C., & Malone, K. M. (2010). Clinical anxiety, cortisol and interleukin-6: Evidence for specificity in emotion–biology relationships. Brain. *Behavior, and Immunity*, *24*, 1074–1077. doi:10.1016/j.bbi.2010.03.003.

Ojala, M., & Garriga, G. C. (2010). Permutation tests for studying classifier performance. *Journal of Machine Learning Research*, *11*, 1833–1863.

Olson, R. S., Cava, W. L., Mustahsan, Z., Varik, A., & Moore, J. H. (2018). Data-driven advice for applying machine learning to bioinformatics problems. *Pacific Symposium on Biocomputing*, *23*, 192–203. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/29218881%0Ahttp://www. pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5890912.

Ormel, J., Jeronimus, B. F., Kotov, R., Riese, H., Bos, E. H., Hankin, B., ... Oldehinkel, A. J. (2013). Neuroticism and common mental disorders: Meaning and utility of a complex relationship. *Clinical Psychology Review*, *33*(5), 686–697. doi:10.1016/j.cpr.2013.04.003.

Pantazatos, S. P., Talati, A., Schneier, F. R., & Hirsch, J. (2014). Reduced anterior temporal and hippocampal functional connectivity during face processing discriminates individuals with social anxiety disorder from healthy controls and panic disorder, and increases following treatment. *Neuropsychopharmacology*, *39*(2), 425–434. doi:10.1038/npp.2013.211.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830. doi:10.1007/s13398-014-0173-7.2. Penninx, B. W. J. H., Beekman, A. T. F., Smit, J. H., Zitman, F. G., Nolen, W.A., Spinhoven, P., ... Van Dyck, R. (2008). The Netherlands Study of Depression and Anxiety (NESDA): Rationale, objectives and methods. *International Journal of Methods in Psychiatric Research*, *17*(3), 121–140. doi:10.1002/mpr.256.

Perez-Cornago, A., Ramírez, M. J., Zulet, MÁ, & Martinez, J. A. (2014). Effectof dietary restriction on peripheral monoamines and anxiety symptoms in obese subjects with metabolic syndrome. *Psychoneuroendocrinology*, *47*, 98– 106. doi:10.1016/j.psyneuen.2014.05.003.

Pitsavos, C., Panagiotakos, D. B., Papageorgiou, C., Tsetsekou, E., Soldatos, C., & Stefanadis, C. (2006). Anxiety in relation to inflammation and coagulation markers, among healthy adults: The ATTICA study. *Atherosclerosis*, *185*(2), 320–326. doi:10.1016/j.atherosclerosis.2005.06.001.

Randall, J. R., Sareen, J., Chateau, D., & Bolton, J. M. (2019). Predicting futuresuicide: Clinician opinion *v.* a standardized assessment tool. *Suicide and Life-Threatening Behavior*, *49*(4), 941–951. doi:10.1111/sltb.12481.

Reece, A. G., & Danforth, C. M. (2017). Instagram photos reveal predictive markers of depression. *EPJ Data Science*, *6*(1), 1–16. 10.1140/epjds/ s13688-017-0110-z.

Rodriguez, B. F., Weisberg, R. B., Pagano, M. E., Bruce, S. E., Spencer, M. A., Culpepper, L., & Keller, M. B. (2006). Characteristics and predictors of fulland partial recovery from generalized anxiety disorder in primary care patients. *Journal of Nervous and Mental Disease*, *194*(2), 91–97.doi:10.1097/01.nmd.0000198140.02154.32.

Schiefelbein, V. L., & Susman, E. J. (2006). Cortisol levels and longitudinal cortisol change as predictors of anxiety in adolescents. *Journal of EarlyAdolescence*, *26*(4), 397–413. doi:10.1177/0272431606291943.

Scholten, W. D., Batelaan, N. M., Penninx, B. W. J. H., Balkom, A. J. L. M., VanSmit, J. H., Schoevers, R. A., & Van Oppen, P. (2016). Diagnostic instability of recurrence and the impact on recurrence rates in depressive and anxiety disorders. *Journal of Affective Disorders*, *195*, 185–190. doi:10.1016/ j.jad.2016.02.025.

Scholten, W. D., Batelaan, N. M., van Balkom, A. J., Penninx, B. W., Smit, J. H., & Van Oppen, P. (2013). Recurrence of anxiety disorders and its predictors. *Journal of Affective Disorders*, *147*(1–3), 180–185. doi:10.1016/ j.jad.2012.10.031.

Spinhoven, P., Batelaan, N. M., Rhebergen, D., van Balkom, A. L., Schoevers, R., & Penninx, B. W. (2016). Prediction of 6-yr symptom course trajectories of anxiety disorders by diagnostic, clinical and psychological variables. *Journal of Anxiety Disorders*, *44*, 92–101. doi:10.1016/j.janxdis. 2016.10.011.

Steyerberg, E. W. (2009). *Clinical prediction models: A practical approach to development, validation, and updating*. New York, NY, US: Springer.

Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformaticsoot*, *8*(1), 1–21. doi:25\r10.1186/1471-2105-8-25.

van Beljouw, I. M., Verhaak, P. F., Cuijpers, P., van Marwijk, H. W., & Penninx, B. W. (2010). The course of untreated anxiety and depression, and determinants of poor one-year outcome: A one-year cohort study. *BMC Psychiatry*, *10*, 86. doi:10.1186/1471-244X-10-86.

Verduijn, J., Verhoeven, J. E., Milaneschi, Y., Schoevers, R. A., van Hemert, A. M., Beekman, A. T. F., & Penninx, B. W. J. H. (2017). Reconsidering the prognosis of major depressive disorder across diagnostic boundaries: Full recovery is the exception rather than the rule. *BMC Medicine*, *15*(1), 1–9. doi:10.1186/s12916-017-0972-8.

Vogelzangs, N., Beekman, A. T. F., De Jonge, P., & Penninx, B. W. J. H. (2013). Anxiety disorders and inflammation in a large adult cohort. *Translational Psychiatry*, *3*(e249), 1–8. 10.1038/tp.2013.27.

Vollebergh, W. A. M., Iedema, J., Bijl, R. V., de Graaf, R., Smit, F., & Ormel, J. (2001). The structure and stability of common mental disorders. *Archives ofGeneral Psychiatry*, *58*(6), 597. doi:10.1001/archpsyc.58.6.597.

Whitfield-Gabrieli, S., Ghosh, S. S., Nieto-Castanon, A., Saygin, Z., Doehrmann, O., Chai, X. J., ... Gabrieli, J. D. E. (2016). Brain connectomics predict response to treatment in social anxiety disorder. *Molecular Psychiatry*, *21*(5), 680–685. doi:10.1038/mp.2015.109.

Wittchen, H. (1994). Reliability and validity studies of the WHO-composite international diagnostic interview (CIDI): A critical review. *Journal of Psychiatric Research*, *28*, 57–84.

Wittchen, H. U., Kessler, R. C., Pfister, H., & Lieb, M. (2000). Why do people with anxiety disorders become depressed? A prospectivelongitudinal community study. *Acta Psychiatrica Scandinavica. 102* (Suppl406), 14–23.

Woo, C.-W., Chang, L. J., Lindquist, M. A., & Wager, T. D. (2017). Building better biomarkers: Brain models in translational neuroimaging. *Nature Neuroscience*, *20*(3), 365–377. doi:10.1038/nn.4478.

World Health Organization (1998). *World health organization, composite international diagnostic interview* (*CIDI*), *core version 2.1*. Geneva: WorldHealth Organization.

Zhao, N., Zhang, Z., Wang, Y., Wang, J., Li, B., Zhu, T., & Xiang, Y. (2019). Seeyour mental state from your walk: Recognizing anxiety and depression through Kinect-recorded gait data. *PLoS ONE*, *14*(5), 1–13. doi:10.1371/ journal.pone.0216591.

Zoccola, P. M., Dickerson, S. S., & Yim, I. S. (2011). Trait and state perseverativecognition and the cortisol awakening response. *Psychoneuroendocrinology*, *36*(4), 592–595. doi:10.1016/j.psyneuen.2010.10.004.

Supplementary materials

Supplementary Methods

Full descriptions of clinical domain measurement instruments

Lifetime history for Anxiety and Depressive Disorders and presence of DSM-IV Panic Disorder, Panic Disorder with Agoraphobia, Agoraphobia, Generalized Anxiety Disorder, Social Anxiety Disorder, MDD, Dysthymia and Alcohol Dependency (one month, six-month, one year recency) at baseline were assessed using the WHO-Composite International Diagnostic Interview (CIDI, version 2.1). The CIDI is a structured interview with good reliability and validity (World Health Organization 1998).

Sleep quality during the past four weeks was assessed with the Women's Health Initiative Insomnia Rating Scale (IRS) (Levine *et al.* 2003). This questionnaire consists of six items and a summary score.

Mood and anxiety symptoms during the past week were assessed with the Mood and Anxiety Scoring Questionnaire (Watson *et al.* 1995). The MASQ consists of 30 individual items and three summary scores: positive affect, negative affect and levels of somatization.

Presence of lifetime bipolar symptoms were assessed with the Mood Disorder Questionnaire (MDQ) (Hirschfeld *et al.* 2000). This questionnaire consists of 13 individual items and a summary score.

Levels of general distress and somatization during the past week were measured with the Four-Dimensional Symptom Questionnaire (4DSQ) (Terluin 1996). These two dimensions were measured in 32 items and two summary scores.

Levels of pathological worrying tendencies were assessed using the 11-item self-report version of the Penn State Worry Questionnaire (PSWQ) (Drost *et al.* 2012).

Depressive symptoms during the past week were assessed with the Inventory of Depressive Symptomatology-SR (IDS-SR) (Rush *et al.* 1986, 1996). The IDS-SR consists of 30 individual items and a summary score.

Symptoms of anxiety during the past month were assessed with the Beck Anxiety Inventory (BAI), a 21-item self-report questionnaire (Beck *et al.* 1988).

Presence of avoidance behaviors was assessed with the Fear Questionnaire (FQ), a 15-item self-report questionnaire (Marks & Mathews 1979).

Presence of childhood trauma was assessed with a 3-item adaptation from the NEMESIS questionnaire (Wiersma *et al.* 2009).

Symptoms of suicidality during the past week were assessed with the 5 items from the Suicidal Ideation Scale (SSI) that refer to the current state (Beck *et al.* 1979).

The total number of negative life-events during the past year was assessed with the Brugha questionnaire (Brugha *et al.* 1985).

Duration of anxiety symptoms during the last four years prior to baseline were assessed with the Life Chart Interview (LCI), a structured retrospective interview using a calendar approach (Lyketsos *et al.* 1994). The LCI has adequate reliability and validity (Warshaw *et al.* 1994). Due to large proportions of missingness, time spent with depressive and avoidance symptoms were not included in analyses.

Subjects' convictions about the importance of care and their past experiences with care in relation to mental health problems were assessed with the 36-item QUality Of care Through the Eyes of the patient (QUOTE): Anxiety/ Depression version (Sixma *et al.* 1998).

Subjects' perceived need for care at various domains during the past six months was assessed with the 20-item Perceived Need for Care Questionnaire (PNCQ) (Meadows *et al.* 2000). After missing data handling, 14 items were included in analyses.

Levels of disability were measured with the WHO-Disability Assessment Schedule (WHO-DAS II), a 36-item self-report questionnaire measuring levels of disability (Chwastiak & Von Korff 2003). Four work-related items were omitted from further analyses due to large proportions of missing values.

All baseline pharmacotherapeutic use was assessed using inspection of medication containers and coded according to Anatomical Therapeutic Chemical (ATC) codes at baseline (WHO Collaborating Centre for Drug Statistics Methodology n.d.). Historic use of psychopharmacotherapeutics during the three-year period prior to baseline was reported retrospectively.

Full descriptions of psychological domain measurement instruments

Anxiety sensitivity was measured with the 16-item Anxiety Sensitivity Index (ASI) (Reiss *et al.* 1986).

Cognitive reactivity to sadness was measured with the 34-item Leiden Index of Depression Sensitivity (LEIDS) (Van der Does 2002).

Levels of mastery were assessed with an adapted version of the Pearlin Mastery Scale, consisting of 5 items (Pearlin & Schooler 1978; Kempen *et al.* 1998).

Personality structure was assessed with the NEO Five-Factor Inventory, a shortened version of the Revised NEO Personality Inventory (McCrae & Costa 2004). In our analyses, all 60 individual items, as well as domain scores from the five domains were used.

In order to be able to assess relative levels of these psychological traits, summary scores were standardized at the level of the whole NESDA sample.

Full descriptions of sociodemographic domain measurement instruments

Information on sociodemographic information from subject was gathered in a structured manner by face-to-face interviews with trained research assistants. Information gained referred to demographic characteristics (6 items), employment status (6 items), marital status (2 items, of which one was omitted due to high number of missing values), sexual preference (2 items, of which one was omitted due to high number of missing values), housing status (5 items), sources and level of income (11 items), religion status (3 items of which two were omitted due to high number of missing values), family and household decomposition (23 items, of which seventeen were omitted due to high number of missing values), and participation in various leisure activities (23 items, of which three were omitted due missing data). Employment status was analyzed categorically, but presented in the

descriptive statistics table dichotomously. Categories that we presented as 'currently employed' included 'now employed', 'self employed' and 'on pregnancy or maternity leave'. The remaining categories were 'occupationally disabled', 'on sickness benefit', 'early retirement', 'unemployed' and 'other'.

Current levels of loneliness were assessed with the de Jong-Gierveld loneliness scale, an 11 item self-report questionnaire (de Jong-Gierveld & Kamphuls 1985). Severe loneliness was defined as a maximum score of 11.

Levels of current social support were assessed with the 38-item Close Person Inventory (CPI) (Stansfeld & Marmot 1992). However, due to large proportions of missing data, only 3 items were included in our analyses.

Full descriptions of biological domain measurement instruments

The number of chronic diseases with or without treatment were assessed using a 21-item face-to-face interview (Penninx *et al.* 2008). Subjects were asked for presence of 30 common chronic somatic diseases and were able to report any additional diseases they may have. This yielded 2 items that were both included.

Levels of chronic pain during the past 6 months were assessed with the chronic graded pain scale in which levels of chronic pain are summarized in a single ordinal item consisting of 5 grades of pain (Von Korff *et al.* 1992).

The current menstrual cycle status was assessed in five self-reported items, of which 4 were used in analyses.

Body Mass Index (BMI) and hip/waist circumference ratio were measured by a trained research assistant, in accordance to international standards. (World Health Organisation 1989)

Autonomic nervous system function was reflected by measurements of mean heart rate, heart rate variability, inter-beat-interval, pre-ejection period, aggregated respiratory sinus arrhythmia, and aggregated respiration rate. These were measured during the baseline data collection interview with the Vrije Universiteit Ambulatory Measuring System (Vu-AMS) (de Geus *et al.* 1995).

Systolic and diastolic blood pressure was measured with the OMRON M4 IntelliSense digital blood pressure monitor (*HEM-752A, Omron Healthcare, Inc., Bannockburn, Illinois, USA* n.d.). In NESDA, the average of two measurements was used.

Handedness was assessed by self-report. Hand-grip strength, a proxy for overall muscle strength, was assessed twice with the Jamar dynamometer (Bellace *et al.* 2000; Ashton & Myers 2004).

Whether subjects had a fever or a cold during the last week was assessed by self-report.

Fasting blood samples of NESDA participants were obtained in the morning around 8 am and kept frozen at -80°C. Various laboratory tests were performed on these samples.

Brain-Derived Neurotropic Factor (BDNF), Triglycerides, High Density Cholesterol (HDL), glucose, tryptophan, kynurenine, 3-Hydroxykynurenine, Cystatin C, Urea, Uric acid, Creatinin, Cotinine, Parathyroid hormone (PTH), 25-hydroxy vitamin D, Dehydroepiandrosterone, Dehydroepiandrosterone sulfate (DHEA-S), Sex hormone-binding globulin (SHBG), Estradiol (E2), and Testosterone (nmol/l) were assayed at the Clinical Chemistry department of the VU University Medical Center using standard laboratory procedures. Dehydroepiandrosterone measurements were omitted due to missings.

High-sensitivity plasma levels of C-reactive protein (CRP) were measured in duplicate by an in-house ELISA based on purified protein and polyclonal anti-CRP antibodies.(*Dako, Glostrup, Denmark* n.d.) The CRP assay was standardized against the CRM 470 reference agent. The lower detection limit of CRP was 0.1 mg/l and the sensitivity was 0.05 mg/l.

Plasma Interleukine-6 (IL-6) levels were measured in duplicate by a high sensitivity enzyme-linked immunosorbent assay.(*PeliKine CompactTM ELISA, Sanquin, Amsterdam, The Netherlands* n.d.) The IL-6 assay was standardized against a recombinant human IL-6 standard. The lower detection limit of IL-6 was 0.35 pg/ml and the sensitivity 0.10 pg/ml.

Plasma Tumour necrosis factor-alpha (TNF-α) levels were assayed in duplicate at Good Biomarker Science, Leiden, The Netherlands, using a high-sensitivity solid phase ELISA.(*Quantikine*® HS Human TNF- α Immunoassay, R&D systems Inc, Minneapolis, MN, United States n.d.) The TNF-α assay was calibrated against a highly purified E. coli-expressed recombinant human TNF-α. The lower detection limit of TNF-α was 0.10 pg/ml and the sensitivity 0.11 pg/ml.

Full descriptions of lifestyle domain measurement instruments

Smoking status was assessed with three items from the Fagerström Test for Nicotine Dependence (Heatherton *et al.* 1991). Subjects were divided into current smokers, former smokers and subjects who never smoked. Two items were omitted due to missings.

Number of different psychoactive banned substances used by subjects was assessed by self-report.

The amount of alcohol consumption during the past year was assessed with the Alcohol Use Disorders Identification Test (AUDIT) (Saunders *et al.* 1993).

The levels of physical exercise, expressed in metabolic equivalent of task (MET)-minutes/week during the past week were assessed with the 4-item International physical activity questionnaire (Ainsworth *et al.* 2000).

Extended description random forest classifier

Each Random Forest classifier (RFC) was build using 1000 classification trees (Breiman *et al.* 1984) and the number of randomly selected variables per node was set at the square root of the number of variables (default value). Subsample aggregating (subagging) was used instead of bootstrap aggregating (bagging) to create new random subsets of data points per tree. Subagging allowed for balancing the data set (Chen *et al.* 2004) by sampling the same number of subjects for each class, and improving the validity of variable importance calculations (Strobl *et al.* 2007). The balancing of the classes can improve classification performance in data sets with imbalanced distribution of classes where a classifier might focus on only correctly predicting the majority class by assigning all data to this class. 63.2% of all subjects was used as the subsampling factor. This corresponds to the number of unique subjects in a bootstrap sample when using bagging and is recommended as a default (Boulesteix *et al.* 2012).

Variable importance calculation

The standard calculation of variable importance for RFC has been shown to be biased towards continuous variables and categorical variables with many categories (Strobl *et al.* 2007). To ensure the validity of variable importance calculations it was suggested to use subagging and permutation-based variable importance calculations (Strobl *et al.* 2007; Altmann *et al.* 2010; Hapfelmeier & Ulm 2013). To implement permutation-based variable importance calculations we permuted each variable separately a 1000 times and assessed its variable importance under permutation (Ojala & Garriga 2010). The computed null-distribution was then used to calculate a *P*-value of the actually observed variable importance for each variable.

Difference in selected variables

To compare whether variable importance differed between the two classification tasks the following analysis was conducted: 1. based on the *P*-values calculated from permutation-based variable importance we computed a rank from most important (smallest *P*-value) to least important variable per cross-validation iteration, 2. we averaged this ranking across the cross-validation runs to obtain an average rank for each variable, 3. we calculated the absolute difference of ranks between the two classification tasks, 4. we explored each rank difference which was higher (lower) than the mean rank difference +(-) twice the standard deviation. In this way we could determine which variables average rank changed strongly between the two different classification tasks.

Supplementary Table 1. Baseline characteristics of anxiety disorder sample, group comparisons between patients who had no common mental disorders (n = 362) at two-year follow-up and patients who did have any common mental disorder at follow-up (n = 525)

Two-year common mental disorders status Baseline characteristics	Recovered (n = 362)	Persistent (n = 525)	Statistics	p
Clinical domain				
PD diagnosis	135 (37.3%)	233 (44.4%)	$X^2 = 4.34$	**0.035**
Agoraphobia diagnosis	111 (30.7%)	206 (39.2%)	$X^2 = 6.86$	**0.009**
SAD diagnosis	147 (40.6%)	261 (49.7%)	$X^2 = 7.15$	**0.007**
GAD diagnosis	92 (25.4%)	185 (35.2%)	$X^2 = 9.63$	**0.002**
MDD diagnosis	103 (28.5%)	259 (49.3%)	$X^2 = 38.7$	**<0.001**
Dysthymia diagnosis	26 (7.2%)	120 (22.9%)	$X^2 = 38.3$	**<0.001**
Use of psychotropic medication, *current*	252 (69.6%)	387 (73.7%)	$X^2 = 1.79$	0.181
Avoidance behaviour severity, *mean FQ, current*	30.96 ± 18.64	39.33 ± 20.06	t = -6.37	**<0.001**
Pathological worrying severity, *mean PSWQ, current*	35.05 ± 9.81	39.35 ± 9.47	t = -6.49	**<0.001**
Suicidal thoughts, *SSI, past week*	38 (10.5%)	145 (27.6%)	$X^2 = 38.6$	**<0.001**
Level of distress, *mean 4DSQ, past week*	14.78 ± 8.79	19.83 ± 8.66	t = 8.47	**<0.001**
Depressive symptoms severity, *mean IDS-SR, past week*	24.96 ± 11.82	32.47 ± 12.92	t = -9.14	**<0.001**
Sleep disturbances, *mean ISR, past four weeks*	9.05 ± 4.94	10.23 ± 5.33	t = -3.38	**0.001**
Anxiety symptoms severity, *mean BAI, past month*	15.57 ± 9.40	20.19 ± 10.79	t = -6.77	**<0.001**
Percentage of time spent with anxiety symptoms, *LCI, past four years*	42.4% ± 32.9	52.7% ± 34.2	t = -4.36	**<0.001**
History of childhood life events[1]	68 (18.8%)	95 (18.1%)	$X^2 = 0.07$	0.79
History of childhood trauma[2]	182 (50.3%)	323 (61.6%)	$X^2 = 11.3$	**0.001**
History of serious suicide attempts	43 (11.9%)	110 (21.0%)	$X^2 = 12.1$	**<0.001**
Psychological domain				
Neuroticism, *mean NEO-FF subscale*	39.77 ± 6.88	43.39 ± 6.69	t = -7.78	**<0.001**
Extraversion, *mean NEO-FFI subscale*	35.40 ± 6.44	32.40 ± 6.67	t = 6.93	**<0.001**
Conscientiousness, *mean NEO-FFI subscale*	41.54 ± 6.19	39.16 ± 6.46	t = 5.52	**<0.001**
Agreeableness, *mean NEO-FFI subscale*	43.84 ± 5.13	42.46 ± 5.41	t = 3.87	**<0.001**
Openness, *mean NEO-FFI subscale*	38.13 ± 5.99	38.18 ± 6.28	t = -0.13	0.90
Cognitive reactivity to sadness, *mean LEIDS*	38.77 ± 17.91	46.77 ± 17.71	t = -6.55	**<0.001**
Anxiety sensitivity, *mean ASI*	33.45 ± 9.63	36.02 ± 10.16	t = -3.77	**<0.001**
Mastery, *mean Mastery scale*	16.49 ± 3.83	13.82 ± 4.00	t = 9.97	**<0.001**
Sociodemographic domain				
Age in years	40.94 ± 12.25	42.59 ± 12.13	t = -1.97	**0.047**
Education years	**12.16** ± 3.29	**11.68** ± 3.37	**t = 2.12**	**0.034**
Female gender	**252 (69.6%)**	353 (67.2%)	$X^2 = 0.56$	0.45
Currently employed	218 (60.2%)	268 (51.0%)	$X^2 = 7.28$	**0.007**
Has children	168 (46.4%)	239 (45.5%)	$X^2 = 0.07$	0.80
Current severe loneliness	29 (8.0%)	76 (14.5%)	$X^2 = 8.50$	**0.004**

Biological domain				
Number of chronic somatic diseases	0.61 ± 0.84	0.75 ± 0.96	t = -2.25	**0.025**
Chronic pain with high disability	67 (18.5%)	154 (29.3%)	X^2 = 13.4	**<0.001**
BMI	25.28 ± 4.64	25.77 ± 5.38	t = -1.47	0.14
Mean heart rate (bpm)	72.03 ± 9.76	71.75 ± 9.89	t = 0.40	0.69
Systolic blood pressure (mmHg)	135.7 ± 20.91	136.5 ± 18.4	t = -0.62	0.53
CRP (mg/L, n=876)	2.70 ± 4.10	2.99 ± 5.83	t = -0.88	0.38
IL-6 (pg/ml, n=876)	1.23 ± 3.05	1.43 ± 3.08	t = -0.94	0.35
TNF-α (pg/ml, n=871)	1.09 ± 1.36	1.04 ± 1.09	t = 0.55	0.58
BDNF(ng/ml, n=865)	9.26 ± 3.57	9.14 ± 3.54	t = 0.48	0.63
Lifestyle domain				
Former smoker	119 (32.9%)	153 (29.1%)	X^2 = 3.99	0.14
Current smoker	125 (34.5%)	216 (41.1%)		
Low physical activity, *past week*	64 (18.9%)	137 (27.3%)	X^2 = 11.5	**0.003**
High physical activity, *past week*	129 (38.1%)	144 (28.7%)		
Any substance use, *past week*	22 (6.1%)	44 (8.4%)	X^2 = 1.65	0.20
Hazardous drinking or alcohol dependency[3] *past year*	76 (21.1%)	120 (22.9%)	X^2 = 0.38	0.54

PD: Panic Disorder; SAD: Social Anxiety Disorder; GAD: Generalized Anxiety Disorder; MDD: Major Depressive Disorder; FQ: Fear Questionnaire; PSWQ: Penn State Worry Questionnaire; SSI: Suicidal Ideation Scale; 4DSQ: Four Dimensional Symptom Questionnaire; IDS-SR: Inventory of Depressive Symptomatology-SR; ISR: Insomnia Rating Scale; BAI: Beck's Anxiety Inventory; LCI: Life chart interview; NEO-FFI: NEO Five-Factor Inventory; LEIDS: Leiden Index of Depression Sensitivity; ASI: Anxiety Sensitivity Index; BMI: Body Mass Index; CRP: c-reactive protein; IL-6: interleukin-6; TNF-α: tumor necrosis factor-α; BDNF: Brain Derived Neurotrophic Factor.

[1] childhood life events (<16 years of age) were parental divorce, being placed in a juvenile prison, raised in a foster family, placed in a child home, death of a parent.

[2] childhood trauma included emotional neglect, psychological abuse, physical abuse and sexual abuse

[3] as measured with the AUDIT. Scores above 8 are reflective of hazardous drinking, scores at 13 or higher (females) and 15 or higher (males) are indicative of probable alcohol dependency.

Supplementary Table 2: Consistently selected significant variables in the recovery from anxiety disorders classification

Item	Description	Selection Frequency [%]
NEO-FFI item 31	I rarely feel fearful or anxious	98
IDS-SR item 27	Panic/Phobic symptoms	98
WHO DAS item 38	How much embarrassment did you experience because of your health problems during the past 30 days?	97
FQ item 05	Walking alone in a busy street	95
CIDI PDA 1m	Panic with agoraphobia - past month	94
Life chart item 01	percent of time with anxiety symptoms	87
WHO DAS item 16	Dealing with people you do not know?	83
NEO-FFI, neuroticism	Anxiety alternative rationally derived decomposition of neuroticism domain	78
CIDI PDA 12m	Panic with agoraphobia - past year	77
IDS-SR item 07	Feeling Anxious or Tense	74
4DSQ item 05	During the past week did you feel: tense?	70
NEO-FFI, item 22	I like to be where the action is	69
MASTERY item 04	I often feel helpless dealing with the problems of life	66
CIDI SAD l	Social Anxiety Disorder - in lifetime	63
CIDI PDA 6m	Panic with agorafobia - past 6 months	61
BAI item 19	Faint, lightheaded	59
4DSQ item 02	During the past week did you suffer from: worry?	54

Supplementary Table 3: Consistently selected significant variables in the recovery from common mental disorders classification

Variable Name	Description	Selection Frequency [%]
MASQ item 03	Felt successful	100
4DSQ item 01	During the past week did you suffer from: feeling down or depressed?	100
CIDI dysthymia 1m	Dysthymia - past month	100
IDS-SR item 05	Feeling Sad	100
4DSQ item 11	During the past week did you feel: that you can't enjoy anything anymore?	100
CIDI dysthymia 6m	Dysthymia - past 6 months	100
4DSQ item 09	During the past week did you feel: that you can't cope anymore?	100
4DSQ item 08	During the past week did you feel: that you can no longer take interest in the people and things around you?	100
4DSQ item 02	During the past week did you suffer from: worry?	100
4DSQ item 10	During the past week did you feel: that you can't face it anymore?	100
Mastery item 02	Some of my problems I cannot seem to solve at all	100
NEO-FFI, neuroticism	Selfreproach	100
Mastery item 04	I often feel helpless dealing with the problems of life	98
Mastery item 05	Sometimes I feel like a play ball of life	98
4DSQ item 05	During the past week did you feel: tense?	97
CIDI MDD 12m	Major Depression - past year	97
CIDI dysthymia 12m	Dysthymia - past year	96
SSI item 02	Desire to die	95
IDS-SR item 08	Response of Your Mood to Good or Desired Events	95
Mastery item 01	I have little control about the things that happen to me	95
IDS-SR item 07	Feeling Anxious or Tense	93
CIDI MDD 1m	Major Depression - past month	93
MASQ item 25	Had trouble making decisions	91
IDS-SR item 21	Capacity for Pleasure or Enjoyment (excluding sex)	90
MASQ item 22	Felt really "up" or lively	88
Life chart anxiety	percent of time with anxiety symptoms	87
BAI item 19	Faint, lightheaded	83
MASQ item 06	Felt really happy	82
MASQ item 13	Felt dissatisfied with everything	79
MASQ item 11	Felt like I was having a lot of fun	79
SSI item 01	Desire to live	76
Loneliness item 03	I experience a general sense of emptiness	73
IDS-SR item 12	Quality of Your Mood	72

BAI item 08	Unsteady	70
NEO-FFI, neuroticism	depression	69
IDS-SR item 15	Concentration/Decision Making	68
MASQ item 23	Felt inferior to others	66
MASTERY item 03	There is not much that I can do to change important things in my life	64
MASQ item 29	Felt really good about myself	62
MASQ item 28	Worried a lot about things	62
QUOTE item 10	Started feeling in control over my problems.	60
PSWQ item 02	Many situations make me worry	59
NEO-FFI, conscientiousness	Orderliness	59
QUOTE item 16	The professional taught me how to deal with future symptoms.	54
QUOTE item 05	The general practicioner explained the pros and cons of different medications.	53
MASQ item 21	Was short of breath	53
IDS-SR item 30	Leaden Paralysis/Physical Energy	52
NEO-FFI, neuroticism	Negative affect	51

Supplementary Table 4: Variables which were more (or less) important in the broad perspective (recovery from anxiety and affective disorders) in comparison to the narrow perspective (recovery from anxiety disorders).

Variable Name	Description
	More important for narrow perspective
WHO DAS item 16	Dealing with people you do not know
BAI item 21	Hot, cold sweats
FQ item 04	Traveling alone by train or bus
IDS-SR item 26	Other bodily symptoms
BAI item 05	Fear of worst happening
CIDI SAD l	Social Anxiety Disorder - in lifetime
Blood plasma item 02	Kynurenine (µmol/l)
WHO DAS item 38	How much embarrassment did you experience because of your health problems during the past 30 days?
CIDI SAD 12m	Social Anxiety Disorder - past year
	More important for broad perspective
loneliness item 03	I experience a general sense of emptiness
QUOTE item 16	The professional taught me how to deal with future symptoms
4DSQ item 07	During the past week did you feel: that you just can't do anything anymore?
CIDI MDD 1m	Major Depression - past month
MASQ item 26	Felt like I had a lot of energy
CIDI MDD 6m	Major Depression - past 6 months
MASQ item 10	Felt hopeless
QUOTE item 10	Started feeling in control over my problems.
SSI item 03	Reasons for living or dying
CIDI dysthymia 12m	Dysthymia - past year
MASQ item 13	Felt dissatisfied with everything
CIDI MDD 12m	Major Depression - past year
QUOTE item 17	The professional reduced my symptoms.
4DSQ item 14	During the past week: tingling in the fingers
4DSQ item 06	During the past week did you feel: easily irritated
IDS-SR item 02	Sleep During the Night
4DSQ item 04	During the past week did you suffer from: listlessness?

Supplementary Table 5: Evaluation of the two-year recovery from anxiety disorders classification using a transfer learning approach [mean (SD)]

Domains	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV
Clinical	0.71 (0.05)	62.5 (5.3)	65.4 (8.7)	59.7 (7.3)	0.53 (0.06)	0.72 (0.06)
Psychological	0.67 (0.05)	62.0 (5.0)	61.6 (8.6)	62.4 (6.6)	0.53 (0.05)	0.70 (0.05)
Socio-demographic	0.65 (0.06)	60.6 (5.7)	64.4 (8.7)	56.7 (7.3)	0.51 (0.06)	0.70 (0.06)
Biological	0.57 (0.05)	55.4 (4.7)	57.3 (7.8)	53.6 (6.4)	0.46 (0.04)	0.65 (0.05)
Lifestyle	0.53 (0.06)	51.7 (4.5)	62.0 (7.2)	41.5 (6.9)	0.42 (0.04)	0.61 (0.06)
Combination	0.71 (0.05)	63.3 (4.8)	65.0 (8.5)	61.7 (6.0)	0.54 (0.05)	0.72 (0.05)

AUC, area-under-receiver-operator-curve; PPV, positive predictive value; NPV, negative predictive value

References

Ainsworth BE, Bassett Jr. DR, Strath SJ, Swartz AM, O'Brien WL, Thompson RW, Jones DA, Macera CA, Kimsey CD (2000). Comparison of three methods for measuring the time spent in physical activity. Medicine & Science in Sports & Exercise

Altmann A, Toloși L, Sander O, Lengauer T (2010). Permutation importance: A corrected feature importance measure. Bioinformatics 26, 1340–1347.

Ashton LA, Myers S (2004). Serial Grip Strength Testing- Its Role In Assessment Of Wrist And Hand Disability. The Internet Journal of Surgery 5

Beck A, Epstein N, Brown G, Steer R (1988). An inventory for measuring clinical anxiety: Psychometric properties. Journal of consulting and Clinical Psychology 56, 893–897.

Beck AT, Kovacs M, Weissman A (1979). Assessment of suicidal intention: The Scale for Suicide Ideation. Journal of Consulting and Clinical Psychology

Bellace J V., Healy D, Besser MP, Byron T, Hohman L (2000). Validity of the Dexter Evaluation System's Jamar dynamometer attachment for assessment of hand grip strength in a normal population. Journal of Hand Therapy

Boulesteix AL, Janitza S, Kruppa J, König IR (2012). Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 2, 493–507.

Breiman L, Friedman JH, Olshen RA, Stone CJ (1984). Classification and regression trees. Wadsworth & Brooks/Cole Advanced Books & Software: Monterey, CA.

Brugha T, Bebbington P, Tennant C, Hurry J (1985). The List of Threatening Experiences: a subset of 12 life event categories with considerable long-term contextual threat. Psychological Medicine 15, 189.

Chen C, Liaw A, Breiman L (2004). Using Random Forest to Learn Imbalanced Data. University of California, Berkeley

Chwastiak LA, Von Korff M (2003). Disability in depression and back pain: Evaluation of the World Health Organization Disability Assessment Schedule (WHO DAS II) in a primary care setting. Journal of Clinical Epidemiology 56, 507–514.

Dako, Glostrup, Denmark (n.d.).

Van der Does W (2002). Cognitive reactivity to sad mood: Structure and validity of a new measure. Behaviour Research and Therapy

Drost J, Der Does AJW, Antypa N, Zitman FG, Van Dyck R, Spinhoven P (2012). General, specific and unique cognitive factors involved in anxiety and depressive disorders. Cognitive Therapy and Research 36, 621–633.

de Geus EJC, Willemsen GHM, Klaver CHAM, van Doornen LJP (1995). Ambulatory measurement of respiratory sinus arrhythmia and respiration rate. Biological Psychology 41, 205–227.

Hapfelmeier A, Ulm K (2013). A new variable selection approach using Random Forests. . Elsevier B.V. Computational Statistics and Data Analysis 60, 50–69.

Heatherton TF, Kozlowski LT, Frecker RC, Fagerstrom K-O (1991). The Fagerstrom Test for Nicotine Dependence: a revision of the Fagerstrom Tolerance Questionnaire. Addiction 86, 1119–1127.

HEM-752A, Omron Healthcare, Inc., Bannockburn, Illinois, USA (n.d.).

Hirschfeld RMA, Williams JBW, Spitzer RL, Calabrese JR, Flynn L, Keck J, Lewis L, McElroy SL, Post RM, Rapport DJ, Russell JM, Sachs GS, Zajecka J (2000). Development and validation of a screening instrument for bipolar spectrum disorder: The mood disorder questionnaire. American Journal of Psychiatry

de Jong-Gierveld J, Kamphuls F (1985). The Development of a Rasch-Type Loneliness Scale. Applied Psychological Measurement 9, 289–299.

Kempen GIJ, Brilman EI, Ormel J (1998). The Groningen Longitudinal Aging Study (GLAS) on functional status, well-being and need of care: Objectives, design and preliminary results . Tijdschrift voor Gerontologie en Geriatrie

Von Korff M, Ormel J, Keefe FJ, Dworkin SF (1992). Grading the severity of chronic pain. Pain 50, 133–149.

Levine DW, Kaplan RM, Shumaker SA, Kripke DF, Bowen DJ, Naughton MJ, Lewis MA (2003). Reliability and validity of Women's Health Initiative Insomnia Rating Scale. Psychological Assessment

Lyketsos CG, Nestadt G, Cwi J, Heithoff K, et al (1994). The Life Chart Interview: A standardized method to describe the course of psychopathology. International Journal of Methods in Psychiatric Research 4, 143–155.

Marks IM, Mathews AM (1979). Brief standard self-rating for phobic patients. Behaviour Research and Therapy 17, 263–267.

McCrae RR, Costa PT (2004). A contemplated revision of the NEO Five-Factor Inventory. Personality and Individual Differences 36, 587–596.

Meadows G, Harvey C, Fossey E, Burgess P (2000). Assessing perceived need for mental health care in a community survey: development of the Perceived Need for Care Questionnaire (PNCQ). Social psychiatry and psychiatric epidemiology 35, 427–35.

Ojala M, Garriga GC (2010). Permutation Tests for Studying Classifier Performance. Journal of Machine Learning Research 11, 1833–1863.

Pearlin LI, Schooler C (1978). The Structure of Coping. Journal of Health and Social Behavior 19, 2.

PeliKine CompactTM ELISA, Sanquin, Amsterdam, The Netherlands (n.d.).

Penninx BWJH, Beekman ATF, Smit JH, Zitman FG, Nolen WA, Spinhoven P, Cuijpers P, De Jong PJ, Van Marwijk HWJ, Assendelft WJJ, Van Der Meer K, Verhaak P, Wensing M, De Graaf R, Hoogendijk WJ, Ormel J, Van Dyck R (2008). The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. International journal of methods in psychiatric research 17, 121–40.

Quantikine® HS Human TNF- α Immunoassay, R&D systems Inc, Minneapolis, MN, United States (n.d.).

Reiss S, Peterson RA, Gursky DM, McNally RJ (1986). Anxiety sensitivity, anxiety frequency and the predictions of fearfulness. Behaviour Research and Therapy 24, 1-8 ST-Anxiety sensitivity, anxiety frequency a.

Rush AJ, Giles DE, Schlesser MA, Fulton CL, Weissenburger J, Burns C (1986). The inventory for depressive symptomatology (IDS): Preliminary findings. Psychiatry Research 18, 65–87.

Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH (1996). The Inventory of Depressive Symptomatology (IDS): psychometric properties. Psychological medicine 26, 477–486.

Saunders JB, Aasland OG, Babor TF, De La Fuente JR, Grant M (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption-II. Addiction 88, 791–804.

Sixma HJ, Kerssens JJ, Campen C van, Peters L (1998). Quality of care from the patients' perspective: from theoretical concept to a new measuring instrument. Health Expectations 1, 82–95.

Stansfeld S, Marmot M (1992). Deriving a survey measure of social support: The reliability and validity of the close persons questionnaire. Social Science & Medicine 35, 1027–1035.

Strobl C, Boulesteix A-L, Zeileis A, Hothorn T (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinformaticsoot 8, 1–21.

Terluin B (1996). De Vierdimensionale Klachtenlijst (4DKL) - Een vragenlijst voor het meten van distress, depressie, angst en somatisatie. [The Four-Dimensional Symptom Questionnaire (4DSQ). A questionnaire to measure distress, depression, anxiety, and somatization]. Huisarts en Wetenschap

Warshaw MG, Keller MB, Stout RL (1994). Reliability and validity of the longitudinal interval follow-up evaluation for assessing outcome of anxiety disorders. Journal of Psychiatric Research 28, 531–545.

Watson D, Weber K, Assenheimer JS, Clark L a, Strauss ME, McCormick R a (1995). Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. Journal of abnormal psychology 104, 3–14.

WHO Collaborating Centre for Drug Statistics Methodology (n.d.). ATC/DDD Index 2019

Wiersma JE, Hovens JGFM, Van Oppen P, Giltay EJ, Van Schaik DJF, Beekman ATF, Penninx BWJH (2009). The importance of childhood trauma and childhood life events for chronicity of depression in adults. Journal of Clinical Psychiatry

World Health Organisation (1989). Measuring obesity : classification and description of anthropometric data. World Health Organization Regional Office for Europe: [Copenhagen].

World Health Organization (1998). World Health Organization, Composite International Diagnostic Interview (CIDI), core version 2.1. Geneva.

Chapter 8.

General summary and discussion

The main aims for this thesis were twofold. The first aim was to increase knowledge on factors that negatively impact the clinical course in anxiety disorders. The second aim was to develop and test statistical models for predicting poor clinical course in anxiety disorders over time. Chapters 2, 3, and 4 focussed on the first aim, whereas chapters 5, 6, and 7 focussed on the second aim.

General summary

In **chapter 2** the aim was to implement a screening programme for anxiety disorders in patients who visited the cardiac emergency department (CED) with acute chest pain. 'Non-cardiac chest pain' (NCCP) is diagnosed in patients presenting with acute chest pain when a cardiac cause for the chest pain can be ruled out. From the literature it is clear that over 50% of patients with acute chest pain are diagnosed with NCCP. A substantial proportion (12-41%) of NCCP have a panic disorder (PD). In light of these high prevalence numbers, the aim of this paper was to screen for PD and other psychiatric disorders in NCCP patients. A sample of 252 adult patients who presented with acute chest pain and were diagnosed with NCCP at the CED in the VU-University Medical Center between 2012 and 2013 were included and eligible for screening. The screening programme consisted of two phases: the first phase was performed at the CED by trained CED nurses and consisted of administration of the Hospital Anxiety and Depression Scale (HADS) after the diagnosis of NCCP was made. Patients who scored above a predefined cut-off score were eligible for the second phase of screening. This second phase was scheduled on a later timepoint at the psychiatry department and consisted of a structured interview into different psychiatric diagnoses using the Composite International Diagnostic Interview (CIDI).

Unfortunately, the first phase of screening was initiated in only 60 out of 252 (23.8%) eligible NCCP patients. This was largely due to low staff adherence, who prioritized other tasks over initiating the screening programme. Patient adherence was a lot better in the first phase of screening: 51 out of 60 patients (85.0%) who were offered screening adhered to the administration of the HADS. Nearly half of screened NCCP patients (24/51 or 47.1%) scored above the HADS cut-off and were eligible for the second phase of screening. Out of these, 12 patients (50%) refused further participation and 8 NCCP patients were diagnosed with a psychiatric

disorder, of which two were diagnosed with PD. Eventually, the screening programme was not deemed feasible. The main barrier for implementation was low staff adherence.

In **chapter 3** the aim was to assess relative impact of anxiety-and depressive disorders (ADDs) and chronic somatic diseases (CSDs) on levels of disability and work impairment. Patients with ADDs are burdened with high levels of disability and work impairment. The same goes for patients with CSDs, such as low back pain, migraine, diabetes and obesity. Furthermore, ADDs often co-occur with CSDs. This physical-mental (PM) comorbidity further reduces levels of functioning and is thought to hamper treatment effects for somatic diseases. Likewise, PM-comorbidity likely negatively impacts the course of ADDs. The aim of this study was to cross-sectionally investigate the effects of ADDs, different categories of CSDs, and PM-comorbidity on functional outcomes in adult patients. The sample was derived from the Netherlands Study of Anxiety and Depression (NESDA) and included patients with ADDs and controls (total n=2,371). Anxiety disorders included generalized anxiety disorder (GAD), social anxiety disorder (SAD), agoraphobia and PD. Depressive disorders included major depressive disorder and dysthymia. Presence of 30 different CSDs was assessed. Only somatic diseases for which patients received treatment or medication were counted. In the sample, patients with ADDs more often had any CSD. From adjusted logistic regression analyses an odds ratio of 1.34 (95% confidence interval (CI) 1.09-1.64) was derived for patients with ADDs for having any of the CSDs. The different CSDs were divided into seven categories: respiratory, cardio-metabolic, musculoskeletal, gastrointestinal, neurological, endocrine and cancers. Out of these, only gastrointestinal diseases were significantly more present in patients with ADDs (OR=3.29, 95% CI 2.15-5.05).

From descriptive statistics it was evident that ADDs were associated with worse functional status: the total standardized disability score, measured with the World Health Organisation Disability Assessment Schedule (WHO-DAS II) was 29.0 ± 16.4 whereas the total disability score in controls was 7.8 ± 9.3 ($t=-30.7$, $p<0.001$). Multivariate linear regression analyses showed that ADDs were related to the highest levels of disability (main effect $\beta= 20.1$). CSDs were also associated with disability (main effect $\beta= 3.88$). Interestingly, there was an interaction effect present for CSDs and ADDs on levels of disability. In the interaction model, the regression formula was disability = $c + \beta1*ADDs + \beta2*$ CSDs + $\beta3*ADDs*CSDs$. In

this formula, c was 13.1, β1 was 18.6, β2 was 0.80 (not significant) and β3 was 4.06. This implicates that the effects of CSDs on disability are mostly present in patients who also have ADDs. In this interaction model, having a CSD without having an ADD does not increase the levels of disability. It also implies that the effect of comorbid CSDs and ADDs on disability scores is larger than the sum of separate effects for CSDs and ADDs, i.e., there is synergistic effect modification.

Work impairment analyses were performed in a subset of 1,462 employed respondents. From descriptive statistics, it seemed that work impairment was higher in patients with ADDs: 67.6% of patients had at least some absence from work due to health related reasons (absenteeism), whereas only 32.7% of controls had absence from work (X^2 = 163.9, p<0.001). Likewise, ADD patients were more likely to have reduced work performance (presenteeism) in comparison with controls: 60.5% of patients had any form of presenteeism, whereas only 32.5% of controls had any form of presenteeism (X^2 = 132.2, p<0.001). To assess the individual and combined effect of ADDs and CSDs on work impairment outcomes multinomial regression analyses were performed. The work impairment outcomes were categorized and comparisons were made against the 'healthy' categories of no absenteeism and no presenteeism. It seemed that CSDs were only associated with the most critical work impairment outcomes: extended absenteeism OR= 1.42 (95% CI 1.07-1.88) and impaired work performance OR= 1.42 (95% CI 1.08-1.87). ADDs were associated with all work impairment outcomes: short absenteeism OR= 2.88 (95% CI 2.16-3.84), extended absenteeism OR= 6.64 (95% CI 4.69-9.40), reduced work performance OR= 1.83 (95% CI 1.38-2.43) and impaired work performance OR= 7.51 (95% CI 5.11-11.1). These regression coefficients were adjusted for sociodemographics and show a clear difference between ADDs and CSDs on work impairment outcomes with ADDs effects being much more impactful. Additionally, main effect models were made using four predictor categories for combinations of ADDs and CSDs exposure: controls (no ADDs, no CSDs), pure physical (CSDs without ADDs), pure mental (ADDs without CSDs), and PM- comorbidity (ADDs and CSDs). The pure physical group was not associated with absenteeism but was associated with impaired work performance. The pure mental subgroup was significantly associated with all work impairment outcomes. Finally, PM- comorbidity was consequently associated with the worst work impairment outcomes.

These findings highlight the importance of recognizing and treating psychiatric comorbidity in patients with CSDs and likewise to recognizing and treating somatic comorbidity in patients with ADDs in order to counter long-term disability and reduced (work) functioning.

In **chapter 4** the aim was to evaluate different definitions and diagnostic criteria for treatment resistance in anxiety disorders (TR-AD). Currently, no clear criteria for TR-AD exist. This is problematic, as a large proportion of anxiety disorder patients experience suboptimal treatment results. Various different terms are used in the literature, for instance "treatment resistance", "remaining symptomatic", "refractory", and "nonresponse". These terms are used quite freely and interchangeably but often seemingly without taking regard to the actual definitions for these terms. In order to align these different terms a systematic literature review was performed. All scientific works that included some definition or operationalization for TR-AD were included. From these sources, definitions were gathered and were systematically described with regard to different aspects or criteria that comprised each definition. The aim was to collect and integrate all criteria and aspects for TR-AD that are currently used. A secondary aim was to integrate these criteria into a consensus definition for TR-AD.

The search strategy yielded 13,042 unique records. These were assessed first by two researchers on title and abstract. This resulted in 388 records eligible for full-text screening. From these, 62 studies were included into the data synthesis. The included studies all provided a specific definition for TR-AD, or they provided inclusion criteria for TR-AD patients. The selection of studies consisted of reviews, guidelines, book chapters, trials and cohort studies. The anxiety disorders studied included PD (n=33), GAD (n=34), SAD (n=21), specific phobia (n=5) and anxiety disorders in general (n=5). In order to meet criteria for TR-AD, most studies (85.5%) included a specified minimal number of failed treatments, ranging from one to five failed treatments. Further criteria included failed pharmacologic treatments (93.5%), failed psychotherapy treatments (29.0%), a specified minimal treatment duration (54.8%), a specific response criterium (41.9%), e.g. a greater than 50% reduction in anxiety severity should be present to no longer meet criteria for TR-AD. Moreover, a threshold for anxiety severity was used in nearly half (46.8%) of studies. Finally, some criteria were used sparingly (less than 10%) in TR-AD definitions: minimal duration of anxiety disorder, presence of functional impairments, presence of comorbidity.

From quantitative analyses, it appeared that study quality did not impact the criteria provided: the same criteria were reported among low and high quality studies.

Finally, the criteria from this systematic review were integrated into a new consensus definition for TR-AD using the most prevalent criteria and aspects identified in this review. According to the consensus definition TR-AD is present if there is at least one failed pharmacological treatment using a first-line antidepressant (SSRI/ SNRI). Treatment failure should be defined as a reduction in symptom severity less than 50% after a treatment period of at least 8 weeks. Second, patients should also have a failed psychotherapy trial, using a first-line psychotherapeutic approach (CBT). This psychotherapy failure should also be defined by a symptom severity reduction that does not exceed 50% and the psychotherapy should be provided according to local protocol and should span a duration of at least 8 weeks. Finally, the symptom severity at assessment of TR-AD should be above a specified threshold. The thresholds described differed across anxiety disorders diagnoses and the cut-off values were based on the most frequently used cut-off values. Consistently using this new consensus definition in studies into TR-AD will increase the homogeneity of the studied population and increase generalizability of findings, thereby bolstering knowledge on TR-AD.

In **chapter 5** the aim was to evaluate a generic staging model for psychiatric disorders for use in anxiety disorders. The main goals were to assess the construct- and predictive validity of the staging model. In the paradigm of clinical staging, different stages can be distinguished that reflect increasing levels of disease progression. Each subsequent disease stage will be associated with a less favourable disease course. Theoretically, the different stages in a staging model should also reflect different underlying pathophysiological processes. However, in psychiatry the underlying pathophysiologic characteristics are multifactorial and thereby less adequate for this purpose. Therefore, currently staging models in psychiatry are mainly based on clinical features.

For the purpose of this study, a heuristic staging model, proposed by a highly regarded Australian research group, was adapted for use in anxiety disorders. All 1,305 NESDA subjects with an anxiety disorder at baseline were included in the analyses. A further 1,115 subjects without current anxiety disorder or depressive disorder, but with presence of risk factors for development of an anxiety disorder were included. At baseline, all 2,420

subjects were assigned to a clinical stage. The stages ranged from stage 0 (asymptomatic, at risk subjects) to stage 4B (chronic symptoms with comorbidity). Stage assignment was based on life chart interviews with regard to anxiety disorder duration as well as on severity of symptoms and on presence of psychiatric comorbidity. From descriptive statistics it was apparent that higher stages showed higher proportions of diagnoses at follow-up timepoints. For instance, at two-year follow-up 2.7% of subjects originally in stage 0 had an anxiety disorder versus 68.0% of subjects originally in stage 4B.

Construct validity was assessed by comparing several baseline clinical characteristics that were not used in stage assignment. It was hypothesized that higher clinical stages would be associated with worse clinical parameters at baseline and that these associations would follow a linear trend along each of the stages. This was indeed the case: patients in higher stages had more childhood trauma, a lower age of onset, more current psychiatric treatments, higher anxiety severity, more social and agoraphobic avoidance and higher levels of worrying. Furthermore, patients in higher stages had more depressive symptoms and higher levels of disability. Nonparametric tests for trends across ordered groups were significant for all predefined validators after Bonferroni correction for repeated statistical testing. This implicates that on average, higher clinical stages were associated with worse clinical characteristics at baseline, indicating adequate construct validity.

To evaluate the predictive validity of the model, a number of analyses were performed. First, the clinical staging model was related to presence of DSM-IV diagnoses over time. Two sets of generalized estimating equations (GEE) were performed to evaluate odds for each baseline clinical stage to have either an anxiety disorder or any psychiatric disorder at three subsequent timepoints (2-year, 4-year and 6-year follow-up). It was hypothesized that higher clinical stages showed higher odds for having anxiety disorder diagnoses and psychiatric disorder diagnoses at each follow-up timepoint. It appeared that this was indeed the case: odds ratio (OR) for having any anxiety disorder at 6-year follow-up for stage 4B was 11.8 (95% CI 8.39-16.6) in comparison with stages 0-1B. Likewise, OR for stage 4B for having any psychiatric disorder at 6-year follow-up was 10.7 (95% CI 7.70–15.0) in comparison with stages 0-1B. For all stages, 6-year proportions of psychiatric diagnoses were lower in comparison to 2-year proportions. For the first outcome, presence of anxiety disorders, a more linear trend across

baseline stages became apparent. For presence of any psychiatric disorders, the B stages were at highest risk. Overall the model showed a significant linear trend across all stages. This implicates that risks of having any anxiety disorder at follow-up are higher in each subsequent stage, but changes of having any psychiatric disorder are highest among B stages who already had psychiatric comorbidity at baseline. Subsequently, predictive validity was assessed using a dimensional approach. The baseline clinical stages were related to follow-up measurements of anxiety severity, depression severity and disability using linear mixed models (LMM), thereby allowing to estimate missing data points. This yielded estimated means for each stage at follow-up timepoints. LMM analyses showed that anxiety severity gradually decreased over time for most stages while retaining the ordering between stages. Stages 2B, 3A and 4A showed comparable anxiety severity over time whereas stages 3B and 4B were burdened with the highest levels of anxiety severity, depression severity and disability.

To summarize, this chapter showed the first successful empirical attempt of applying a staging model to a cohort of anxiety disorder patients and controls. The results show that the studied model has adequate predictive and construct validity, thereby providing an evidence-based staging model for use in clinical care in anxiety disorders.

In **chapter 6** the aim was to assess whether each of the aspects of TR-AD derived in a systematic review (chapter 4 in this thesis) were associated with poor outcomes. As chapter 4 yielded a number of criteria for the definition of TR-AD, the hypothesis was that each of these criteria were individually associated with poor outcomes after treatment. A secondary aim in this study was to develop a dimensional measurement instrument using combined information from each of the individual TR-AD criteria. This dimensional instrument reflects the degree of TR-AD, which could be a relevant addition to the dichotomous approach to TR-AD that was applied in chapter 4.

For the purpose of this study, out of a total of 1,305 NESDA subjects with anxiety disorders at baseline, a sample of 679 subjects that reported having received psychiatric treatments between baseline and 2-year follow-up were included. In this sample, each of the criteria for TR-AD that were derived in the systematic review (chapter 4) were assessed at baseline. These baseline TR-AD criteria included number of first-line pharmacologic treatments previous to inclusion, number of second-line pharmacologic treatments previous to inclusion, number of adequate psychotherapeutic treatments

previous to inclusion, levels of anxiety severity, presence of functional impairments, presence of psychiatric comorbidity, and previous duration of anxiety symptoms. From these literature-derived criteria a measurement tool was developed. In order to do this, each of the TR-AD criteria were scored in accordance with the Dutch Measure for quantification of Treatment Resistance in Depression (DM-TRD). The DM-TRD is a dimensional measurement tool that measures the degree of treatment resistance in depression. This way a measurement tool with a potential range of scoring from 2-23 was made.

At baseline, the average score on the measurement tool was 10.8 ± 2.3. Bivariate logistic regression analyses were used to relate individual baseline TR-AD criteria to treatment outcome after two-years by assessing persistence of anxiety disorder diagnoses at two-year follow-up. High symptom severity at baseline was clearly associated with poor outcomes at two-year: OR for persistence of anxiety disorders was 6.48 (95% CI 3.29-12.8) in comparison with subjects with low symptom severity at baseline. High levels of functional impairments were also associated with poor outcomes at two-year follow up: OR=2.90 (95% CI 1.51-5.63). The same positive associations were found for presence of psychiatric comorbidity (OR=1.73, 95% CI 1.25-2.39) and extended duration of anxiety symptoms (OR = 2.79, 95% CI 1.94-4.03 versus short duration). The previous number of first-line pharmacotherapeutic treatments were not associated with outcomes after two year (OR=1.10, 95% CI 0.87-1.39). Nor were the number of previous second-line pharmacotherapeutics (OR= 1.39, 95% CI 0.91-2.12) or the number of previous psychotherapy trials (OR=1.11, 95% CI 0.73-1.69). The dimensional measurement instrument that was developed using these TR-AD criteria was positively associated with poor outcomes at two-year follow-up: OR= 1.29, 95% CI 1.20-1.39, indicating that the odds for persistence of anxiety disorders after two-years are increased with 1.29 for each point increment in the measurement instrument.

Finally, the psychometric properties for this measurement instrument were assessed. Using the Youden-index the most efficient cut-off value for the measurement instrument was calculated at 11 points or higher. Using this cut-off value, the sensitivity for predicting persistence of anxiety disorders during treatment was 0.70, while the specificity was 0.57. From this, the positive predictive value was calculated at 0.68 and the negative predictive value at 0.60. The area under the curve (AUC) was calculated to be 0.66. This performance can be regarded as moderate. However, in the absence

of a gold-standard test for predicting outcomes after treatment in anxiety disorders, it represents the best performance currently available. Therefore, it seems that the newly developed measurement tool for assessing the degree of TR-AD could be beneficial in assessing the level of TR-AD in individual patients. This could have implications for choosing adequate treatment regimens in individual patients. However, this study should be replicated in a different sample to evaluate the generalizability of these findings.

In **chapter 7** the aim was to develop a machine learning prediction model for the longitudinal course in anxiety disorders based on a wide variety of data. Many risk factors were previously linked to anxiety disorders or to the longitudinal course in anxiety disorders. None, however, have predictive properties that warrant using them for risk prediction in clinical care. Previously, these predictors were mainly studied in isolation instead of as a part of a large scale model. Possibly, combining many putative predictors into a single model could result in a model with adequate predictive properties. Machine learning algorithms are especially suited for problems with many different putative factors, as they are able to use data-driven methods for selecting the most relevant factors.

For the purpose of this study, various putative predictors from five predictor domains were selected. These domains included clinical variables, psychological variables, biological variables, sociodemographic variables and lifestyle variables. The hypothesis was that combining many predictors into a single large scale model would result into a model with adequate predictive properties for longitudinal course in anxiety disorders. The current study was performed in a total of 887 anxiety disorder patients from NESDA. The investigated classifications were twofold: first, recovery from anxiety disorders at two-year follow-up. Second, recovery from all common mental disorders (CMDs) at two-year follow-up. Anxiety disorders included GAD, PD, agoraphobia and SAD. CMDs were defined as either an anxiety disorder, major depressive disorder, dysthymia or alcohol dependency. So, for the purpose of the second classification, recovery is present when no disorder from these disorders groups was diagnosed at follow-up. To build the model, Random Forest Classifiers (RFCs) were used. A RFC is built as an ensemble of many different decision trees, which are trained by considering random subsets of variables and patients for each tree. In this study, each RFC consisted of 1,000 decision trees. At baseline 651 individual item level predictors were selected from the five predictor domains (clinical,

psychological, biological, sociodemographics, lifestyle). After removal of items with too many missing values, 569 items remained. RFCs were built using these items and by using a 10x10 cross validation approach on training sets (90% of the subjects) and test sets (the remaining 10% of the subjects). The algorithm was trained on the combination of all predictor domains, but also on each different predictor domain separately.

At two-year follow-up, 484 patients (54.6%) recovered from anxiety disorders and 362 patients (40.8%) did not have any CMD. The performance of RFCs in predicting recovery of anxiety disorders at two-year follow-up was moderate: the area under the receiver operator curve (AUC) ranged from 0.49 (lifestyle) to 0.67 (clinical) in individual predictor domains. The only domains with statistically significant two-year predictions were the clinical domain (AUC=0.67) and the psychological domain (AUC=0.65). When combining predictors from all domains, the AUC was 0.67. With regard to the second outcome measure (recovery from CMDs) it seemed that predictions were somewhat more precise. The AUC for recovery from CMDs ranged from 0.53 (lifestyle) to 0.70 (clinical) in individual domains. In this analysis the clinical domain (AUC=0.70), the psychological domain (AUC=0.67) and the sociodemographic domain (AUC=0.65) yielded significant predictions. The combination of all domains yielded an AUC of 0.70.

Additional analyses were performed to identify individual items that highly contributed to the predictions. This was done to gain more insight into the relative importance of each of the predictors in the longitudinal course in anxiety disorders. In predicting recovery from anxiety disorders, 17 items were selected in over 50% of the RFCs, thereby fulfilling the criterium for consistent selection. These predictors derived from the clinical and psychological domains and were related to anxiety responses, mostly to anxious arousal. In predicting recovery from CMDs, 48 variables were consistently selected. One consistently selected predictor was from the sociodemographic domain, all others from the clinical and psychological domains. In addition to anxiety items, for this classification many mood-related items were consistently selected. When compared to the variables that were consistently selected in predicting recovery from anxiety disorders, it seemed that recovery from CMDs was more dependent on depression-related variables.

Discussion

Methodological strengths and limitations

Several strengths and limitations need to be addressed in order to appraise the findings reported in this thesis. First, some general strengths and limitations on the NESDA cohort will be discussed. After that, the most relevant strengths and limitations of each of the separate chapters will be addressed.

Four of the chapters presented in this paper used data from the **Netherlands Study on Depression and Anxiety** (NESDA) for data analyses. As described in the general introduction, NESDA is a prospective cohort study designed to evaluate the longitudinal course in depression and anxiety. A major strength of this study is the large sample size (n=2,981 at baseline). Furthermore, a very extensive approach to data collection was chosen. This enabled the operationalization of different models from literature within NESDA. For instance, this allowed the operationalization and validation of the generic staging model presented in chapter 5 for use in anxiety disorders without having to omit major criteria that are used in the model. Also, it allowed for using a wide array of putative predictors (N=651) to build the data-driven prediction model presented in chapter 7.

However, using NESDA also had some limitations. First, repeatedly using the same dataset for data analysis has the downside of reduced generalizability across the reported findings. Repeating some of the analyses in an independent sample would be beneficial to increase the generalizability. Still, previous analyses showed that the longitudinal clinical course of subjects in NESDA is comparable with that in other cohort studies.[1] This bolsters the confidence that the findings in this thesis are in fact generalizable beyond NESDA and beyond the Netherlands. Furthermore, by design NESDA has a large time gap between each wave of measurements. The shortest follow-up period between two full waves of data collection used in these papers was two years. From a practical and economical perspective this decision is very understandable. However, this comes at the cost of having a long period of time in which no data on disease status or on environmental stressors is available. At each subsequent wave of data collection, information on the intervals between waves were retrospectively assessed to make up for this loss. These retrospective assessments are suboptimal and this approach hampered some of the analyses presented in this paper. Finally, for the purpose of course prediction, periods of two years represent a very extensive

timespan. Many other studies into course prediction use shorter follow-up periods as predictions are thereby less vulnerable to unknown outside variables affecting the accuracy of predictions. Therefore, the chances of highly accurate course predictions using NESDA data and a minimal follow-up period of two-years were by design reduced.

An important strength of the screening programme presented in **chapter 2** was its implementation in routine clinical care as provided at the CED. If proven successful, this would implicate a high applicability for this programme across CEDs. Currently, the large group of NCCP patients leave the CED with a *diagnosis per exclusionem* (NCCP). This screening programme could improve the diagnostics for this large group of patients without major alterations to the workflow at the CED. The major limitation in this study was the lack of systematic gathering of data on reasons for low adherence across CED staff and patients. For instance, a qualitative study phase as addition to the current study could have shed more insight into the low adherence and might have yielded useful input for future screening programs.

The major strength of the study presented in **chapter 3** is the rigorous assessments of both chronic somatic diseases (CSDs) and anxiety- and depressive disorders (ADDs). This made the current study the largest study of its kind in which a wide variety of CSDs were assessed, as well as a wide range of ADDs. This enabled us to link these to the functional outcomes of disability and work impairment and assess the presence of interaction effects. Furthermore, the large sample size made it possible to evaluate separate groups of CSDs. A limitation in this study was the reliance on self-report evaluation for assessments of CSDs. This is accompanied by the risk of a self-report bias. In order to reduce risks for self-report bias a sensitivity analysis was performed in which more stringent criteria for CSD diagnoses were applied. Applying these stringent criteria did not change the main outcomes, which bolstered the confidence that the initial analyses were accurate.

The main strength of **chapter 4** was the rigorous approach to delineate the concept of treatment resistant anxiety disorders (TR-AD). Although various authors pinpointed the lack of a consensus definition for TR-AD, no one previously performed a systematic review of the available literature. This approach made it possible to derive a consensus definition thereby removing the necessity of authors and scholars of having to choose one definition over

the other. A major limitation of this study was that the process of integrating the literature-based definitions for TR-AD could not be done without some arbitrary choices by the authors. For instance, the decision of basing the consensus definition on the statistical mode for each criterium within the definition was made by the researchers. Furthermore, some criteria that were only sporadically mentioned were removed. These choices are defendable, but still arbitrary to some extent and this could have introduced a small bias.

The main strength of **chapter 5** in which a staging model was adapted for use in anxiety disorders is the use of an validation approach. In this approach, construct validity was assessed by comparing each of the clinical stages to construct validators. After using a number of anxiety-related questionnaires to assign subjects to clinical stages it was still possible to use other baseline questionnaires as construct validators. A limitation in chapter 5 was the transgression that was made away from the original transdiagnostic framework for the staging model. In the original model, less emphasis is laid on different psychiatric disorders but more on the longitudinal course without taking into account which disorder is present. In adopting this model, the aim was to incorporate the transdiagnostic approach by evaluating presence of comorbidity as a profiler and by choosing general psychopathology validators that reflect that longitudinal course in anxiety disorder should indeed be seen in a transdiagnostic framework.

A major strength in **chapter 6** is the use of literature-based criteria for TR-AD as predictors for subsequent course during treatment. This is an improvement over other measurement instruments that assess the degree of treatment resistance in other psychiatric disorders as these measurement instruments were not based on findings from a systematic review. A limitation in this chapter is the retrospective assessment of treatment exposure during the follow-up and using this assessment as an inclusion criterion. This inclusion criterion was necessary in order to be able to study treatment resistance in a subset of patients that received treatments. But by using a retrospective selection criterion this study cannot be seen as fully prospective. This might have introduced a selection bias or a self-report bias. Additionally, some underreporting of certain types of treatments could have occurred as data on treatments were not gathered via care providers but via subjects. This could have introduced some bias towards the null hypothesis with regard to effect of previous treatments.

The main strength of **chapter 7** in which machine learning algorithms were used for course prediction in anxiety disorders was the use of a data driven approach. The aetiology of anxiety disorders is multifactorial and the precise relationship between different aetiologic factors remains largely unknown. Data driven machine learning methods are especially adept in unravelling these kind of scientific problems. Another strength is the use of cross validation methods to improve the external validity and reduce the risk of overfitting the model to the data. Moreover, using features at an individual item level increased the internal validity for building the model as it enabled previously unrecognized associations between individual variables to be used in the predictions. By contrast, if summary scores would have been used, a lot of information and individual variation would have been lost. A limitation in this chapter was that the levels of evidence for including each of the different putative predictors varied substantially. Theoretically, all predictors should be previously related to anxiety disorder longitudinal course in order to be used validly. In reality, all putative predictors were at least related in some way to anxiety disorders, but for a large proportion this was limited to cross-sectional associations. Although the Random Forest classifier is adept in excluding features without any link to the investigated outcome the accuracy of the prediction model could ultimately suffer from inclusion of many unrelated features as in each iteration a number of unrelated features have to be discarded, leaving less choice for adding useful features. Therefore, including a smaller number of features at baseline could have resulted in improved accuracy in the final models.

Factors that impact the clinical course in anxiety disorders.

As is apparent from the introduction from this thesis, the most prominent factors that impact the clinical course in anxiety disorders are clinical factors like anxiety severity, anxiety duration, presence of comorbidity and other clinical aspects. The first aim in this thesis was to expand on the knowledge about factors that impact the clinical course in anxiety disorders. To describe these factors in the remainder of this discussion an ordering across susceptibility markers, diagnostic markers, prognostic markers and predictive markers will be used. This ordering is often used in research into biological markers for psychiatric disorders and can also be used for markers from different domains.[2,3] These marker categories are helpful as they point towards the mechanism of action of each marker. Currently, most research in anxiety disorders focussed on diagnostic markers. For instance, in PD structural changes to certain brain areas and altered serotonin and noradrenergic network activation in patients were found. These biomarkers

are present in patients but less so in controls, thereby constituting a diagnostic marker. More knowledge on each of the markers types could lead to better understanding of the clinical course in anxiety disorders. All findings from this thesis will be discussed in accordance to this ordering of markers.

Susceptibility markers

Susceptibility markers are present in persons who later develop a particular disorder.[3] Perfect susceptibility markers are always present in individuals who later develop a disorder but never in individuals who do not develop this disorder. From previous research it is apparent that several sociodemographic characteristics are overrepresented in anxiety disorder samples. Well established risk factors for development of anxiety disorders include low socio-economic status, female gender, lifetime history and family history of psychiatric disorders and exposure to childhood adversity such as abuse, neglect and parental problems.[4–6] Furthermore, young age could be seen as a risk factor for development of anxiety disorders due to the young average age of onset. The findings in this thesis underscore the previous findings that low socio-economic status, female gender and younger age are related to subsequent development of anxiety disorders. Lifetime history of anxiety disorders and family history of psychiatric disorders were used in chapter 5 to assign subjects to at-risk stages. These risk factors can be regarded as susceptibility markers for future development of anxiety disorders.

Diagnostic markers

Diagnostic markers are markers that distinguish between patients with a particular disorder and those who do not have the disorder.[2] Perfect diagnostic markers are always present in patients but never in controls. In chapter 2 it was investigated whether non-cardiac chest pain (NCCP) can be used as a diagnostic marker for PD. NCCP often presents with squeezing, pressure-like, or burning sensations in the chest and although cardiologists see it as a diagnosis per exclusionem, the presence of NCCP should instigate a further diagnostic work-up. The list of differential diagnoses for NCCP is extensive and includes both somatic illnesses as well as psychological and psychiatric disorders.[7] Even in absence of psychiatric disorders, symptoms of NCCP can be treated effectively with cognitive behavioural therapy.[8] Due to the overlap in symptomatology between NCCP and panic attacks and due to the high prevalence of PD in cohorts of NCCP patients the symptoms of NCCP could be due to PD. A diagnosis of NCCP could thereby constitute a diagnostic marker for PD. Unfortunately, the study presented in chapter 2 suffered from

implementation problems that led to insufficient adherence in CED staff as well as in patients. Therefore it does not seem feasible to improve the diagnostic process of PD (and other psychiatric disorders) by providing this screening program in the CED. Possibly, the presence of NCCP in PD patients might be a susceptibility marker for PD instead of a diagnostic marker. This could be the case if chest pain is the first symptom in the development of a repeating pattern of panic attacks and avoidance behaviours. Another possibility is that chest pain is a prognostic or predictive marker in PD. This could be the case if the presence of chest pain in PD patients was found to be related to the subsequent course in these patients or is predictive of subsequent treatment effects. Unfortunately, these hypotheses could not be tested using the present study and could be the focus for future research.

Subthreshold anxiety symptoms are a likely diagnostic marker for preclinical high risk stages of anxiety disorders as subjects without anxiety disorders who suffer from subthreshold anxiety symptoms are at high risk for development of subsequent anxiety disorders. This group usually contains many individuals who have some of the aforementioned susceptibility markers. In chapter 5, subjects were assigned to at-risk stage 0 if they had any of the susceptibility markers (previous anxiety disorders, family history for psychiatric disorders or childhood trauma) but did not have any current anxiety symptoms. Prodromal stages 1A and 1B were assigned to subjects with these susceptibility markers who also had subthreshold anxiety symptoms. Longitudinal analyses of these groups showed an incremental increased odds for incident anxiety disorders at various follow-up timepoints. At-risk subjects (stage 0) had 2.7% incidence of anxiety disorders at 2-year follow-up, 3.0% incidence of anxiety disorders at 4-year follow-up and 3.1% incidence of anxiety disorders at 6-year follow-up. Incidence levels were markedly higher in stages 1A and 1B: 14.9% and 21.1% at 2-year follow-up, 11.2% and 15.6% at 4-year follow-up and 7.6% and 15.7% at 6-year follow-up respectively. It seemed that a large subset of subjects in prodromal stages subsequently developed an anxiety disorder. The presence of subthreshold anxiety symptoms along with the presence of susceptibility markers therefore seems an adequate diagnostic marker for the prodromal stage. Possibly, targeting groups of children or adolescents who have susceptibility markers and subthreshold anxiety symptoms could result in preventing development of anxiety disorders.[9,10]

The aim in chapter 4 was delineating a useful definition for treatment resistance in anxiety disorders (TR-AD). By analysing and integrating the available literature on this topic a new consensus definition was created for TR-AD. The full definition was provided earlier. As TR-AD refers to chronic and difficult to treat pathology, this new definition for TR-AD could be used as a diagnostic marker for the more advanced stages of anxiety disorders.

In chapter 5 several criteria were used to assign subjects to different stages of anxiety disorders. The main defining criterium used was duration of anxiety and avoidance symptoms. This is in line with the rationale behind staging models in psychiatry in which disease progression is mainly defined by the previous disease course.[11] Using adjusted logistic regression models is was apparent that anxiety/avoidance duration categories (<30% of the previous months; 30<80% of the previous months; >80% of the previous months) were indeed linked to two-year outcomes. This remained true when adjusting for sociodemographics and other predictor variables. Duration of anxiety disorders could therefore be seen as a diagnostic marker for the different stages of anxiety disorders.

Prognostic markers

Prognostic markers provide information about the clinical course in a particular disorder without taking treatment effects into account.[2] Perfect prognostic markers are present in patients who have the highest probabilities for adverse disease-related events or a faster rate of decline but never in patients who are at low risk for these events or this decline. From analyses in chapter 3 it became apparent that an interaction effect on disability and work impairment outcomes is present between ADDs and CSDs. WHO-DAS II disability scores were 4-5 points higher in ADD patients who also have a CSD. Likewise, work impairment outcomes were the least favourable in ADD patients who also had CSDs. These findings point towards using CSDs as a prognostic marker for worse overall course – in terms of functioning - in anxiety disorders.

From the analysis in chapter 5 it became apparent that a younger age of onset was present in the most severe anxiety disorder patients. The average age of onset in stage 4B (severe, persisting, comorbid anxiety disorders) was 19.2 ± 12.4 years whereas the average age of onset in stage 2A (non-comorbid anxiety disorders with a short duration) was 23.9 ± 13.1 years. There was a

statistical significant trend across increasing stages. Younger age of onset can thereby be regarded a prognostic marker for development of chronic, severe anxiety disorders.

In all stages of anxiety disorders, presence of psychiatric comorbidity was clearly associated with markedly worse outcomes at follow-up, for instance adjusted OR for presence of two-year anxiety disorders were 1.49 (95% CI: 1.16-1.92). Presence of comorbidity should be seen as a prognostic marker. This is in line with a large field of research that consistently showed the course of disorders to be worse if comorbid psychiatric disorders are present.[12,13]

Further prognostic markers were identified in chapter 7 in which a machine learning algorithm was used to predict the naturalistic course over a two-year period in anxiety disorder patients. Using a random forest classifier allowed for identification of different important predictor variables by assessing the frequency with which each of the predictors were selected in each iteration of the algorithm. This was done by permuting each variable separately a 1,000 times and assessing its variable importance under permutation.[14] Features that were selected consistently can be considered prognostic markers for the naturalistic course in anxiety disorder patients. This approach led to 17 out of the available 569 items to be consistently selected (>50%) in the random forest classifier on the main outcome measure. These items included statements on traits on anxious arousal and avoidance behaviours (e.g. "I rarely feel fearful or anxious", of generally avoiding "walking alone in a busy street"). Furthermore, items on personality traits neuroticism and extraversion were consistently selected. Moreover, psychological traits of mastery and anxiety sensitivity were often selected in the random forests. Other poor prognostic markers from descriptive statistics were presence of chronic pain, unemployment and low physical activity. These markers were more prevalent in the sample that subsequently did not recover from anxiety disorders, they were, however, not consistently selected (>50%) in the random forest classifiers. The benefit of prognostic markers could lie in risk stratification. Patients with poor prognostic markers should be identified and allocated targeted treatments that could prevent further disorder progression. It should be noted that in the setting of a naturalistic cohort study, a number of subjects received treatments, which makes it more difficult to identify prognostic markers.

Predictive markers

Predictive markers refer to markers that are linked to treatment results in patients with a particular disorder.[2] Perfect predictive markers are present in all patients who show a specific response to treatments or interventions (for instance remission after psychotherapy) but never in patients who do not show this response. In chapter 6 the aim was to unravel predictive markers for anxiety disorder patients who underwent treatments over the course of a two-year period. A dimensional measurement for the degree of TR-AD predicted treatment results after two-years: the odds for persistence of anxiety disorders were increased by 1.29 for each point increment in the measurement instrument. When assessing the separate variables it appeared that treatment outcomes were less beneficial in anxiety disorder patients with a low educational status (OR for persistence of anxiety disorders = 0.95 (95% CI: 0.90-0.99) per education year. Furthermore, higher anxiety severity predicted persistence of anxiety disorders (OR for severe symptoms= 6.48 (95% CI: 3.29-12.8). Moreover, psychiatric comorbidities were more present in the subset of patients that persisted after treatment (OR=1.73, 95% CI: 1.25-5.63). Finally, a longer duration of anxiety and avoidance symptoms predicted persistence of anxiety disorders after treatment, OR for extended previous duration= 2.79 (95% CI: 1.94-4.03). It should be noted that in chapter 6, only clinical variables were assessed as predictive markers.

Integrated stage-specific model

One of the current challenges in scientific research in psychiatry is forming bridges between fundamental research into psychopathology and clinical research. The DSM- classification system remains dominant in the scientific psychiatric debate even as leading authorities acknowledge the need to reconstruct these constructs and start rebuilding a classification system from the bottom up. This approach is advocated in the Research Domain Criteria as proposed by the National Institute of Mental Health.[15] One of the main goals for which the RDoC criteria were called into life is to work towards translation of biological findings from fundamental research into clinical practice. As DSM-classification systems proved inadequate for this goal, staging models might provide a solution. Staging models in psychiatry were developed to describe the progression of psychiatric disorders across the lifespan. In lower stages a-specific symptoms are present, in later stages psychiatric syndromes have developed or even refractory disorders have formed. In staging models, different stages can be linked to susceptibility markers, diagnostic markers, prognostic markers and predictive markers. This approach will likely yield different markers across clinical stages in a

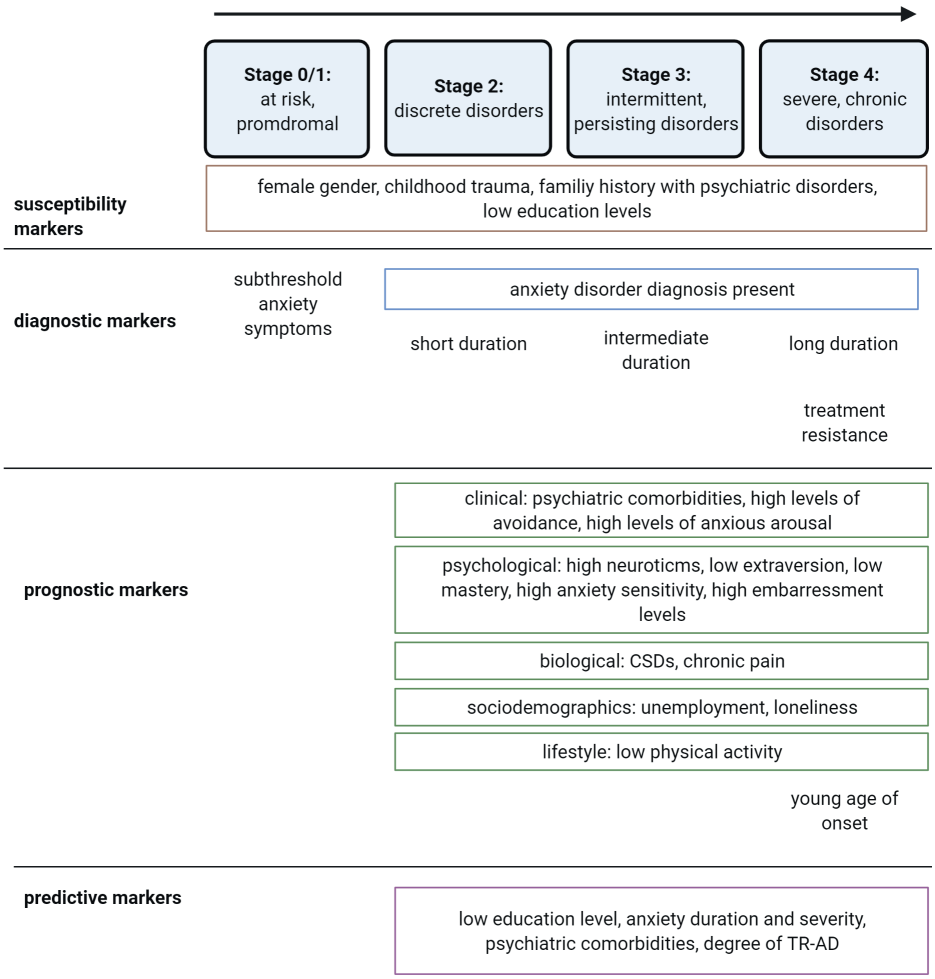


Figure 1. Integrated stage-specific model for anxiety disorders.

way that could not be captured within the framework of DSM-classifications. In this way, staging models could be an important bridge between fundamental research findings and clinical practice. To integrate the main findings of the research in this thesis an integrated stage-specific model was developed. This integrated model for anxiety disorders was modelled to the recent putative biomarker stage-specific model for PD by Cosci and Mansueto.[3] This stage-specific approach has the benefit that biomarkers and clinical markers might be more sensitive, specific and predictive when applied to a specific stage of anxiety disorders instead of to a dichotomous classification.[3] In addition to the model by Cosci, the putative markers were

divided into susceptibility markers, diagnostic markers, prognostic markers and predictive markers as these distinctions are clinically relevant and are vital for development of personalized psychiatry. This model is thereby also in line with the RDoC criteria for future finetuning of classification systems.[15] See figure 1 for the integrated stage-specific model.

Staging in psychiatry

Over the last decade, using clinical staging in psychiatry has become increasingly widespread. The first staging models for psychiatric disorders used an approach for a single disorder or psychiatric syndrome, e.g. for schizophrenia,[16] bipolar disorder,[16] PD[16] and alcohol use disorder.[17] An important benefit of these staging models is expanding the possibilities clinicians have to formalize their individualized assessment of the relative severity of the index disorder.[16] Due to this syndromal approach, use of these models is restricted to health care settings who specialize in providing treatment for certain disorders. More recent studies increasingly advocate using a transdiagnostic approach, e.g. for severe psychotic and mood disorders[18] or for overall psychopathology in adolescents.[19] In these transdiagnostic staging paradigms, prodromal stages of psychiatric disorders are characterized by a-specific clinical features and are not yet distinguishable as specific disorders, while later stages develop into well-defined syndromal psychiatric classifications. In this transdiagnostic approach staging models are described as a "*departure from silo-based diagnostic concepts that populate the current international classification systems*".[19] Comorbidity can be accounted for in transdiagnostic staging models by assessment of all relevant forms of psychopathology.

Anxiety disorders are a group of disorders that seem particularly suitable for assessment via a staging model. Many anxiety disorder patients refrain from seeking treatments for years after development of the first symptoms.[20] As a result, the first episodes likely develop into chronicity.[21,22] Furthermore, anxiety disorders are burdened with high levels of comorbidity with other anxiety disorders or with depressive disorders.[23] Finally, there seems to be substantial diagnostic instability. A little over 30% of persons with an anxiety disorder have a transition to another anxiety disorder within a six-year follow-up period. This proportion increases to 73% in case of a chronic course during follow-up.[24] In this thesis four anxiety disorders were studied in unison. This was done to account for this diagnostic instability but also due to the likelihood of shared aetiology in these disorders. For instance, GWAS findings suggest these anxiety disorders share some genetic aetiology.[25]

Moreover, some debate as to whether staging models should be 'bidirectional' or 'unidirectional' exists. The staging model by McGorry et al applied a unidirectional approach in which clinicians are advised to "note" whether subjects remit spontaneously, or recover after treatment and afterwards, these persons maintain their highest stage they were ever assigned to.[19] A clear rationale for this decision was not provided and their approach has some important limitations. For instance, it leads to a conflation of higher stages, as remitted subjects have lower risks of poor outcomes, not comparable to subjects with current disorders or high symptom severity.[24] Bidirectional models on the other hand are characterized by a design in which subjects can not only progress, but also mitigate across stages. Subjects with remitted disorders are especially crucial in this discussion, as they are the largest group of subjects that would either mitigate (in a bidirectional model) or remain at a higher stage (in a unidirectional model). The model presented in chapter 5 was bidirectional: those with remitted disorders were assigned to stage 0, 1A or 1B, instead of remaining in stages 2, 3, or 4. The clinical course of anxiety disorders can be described as "waxing and waning" disorders.[24,26,27] Therefore, for anxiety disorders, bidirectional staging models might fit best.

However, in contrast to the staging model tailored for use in anxiety disorders presented in chapter 5, some authors argue that staging models should be *fully* transdiagnostic.[19] Instead, the staging model presented in chapter 5 could be termed *half* transdiagnostic as it did incorporate psychiatric comorbidity but did not assess the severity of the comorbidities. Instead, patients with different types of comorbidities were all assigned to the same 'B' substage. This was likely an oversimplification of the impact of psychiatric comorbidity on anxiety disorder course. However, the model presented in chapter 5 was validated in a transdiagnostic way by assessing general psychopathology validators. In this way it became apparent that higher clinical stages showed not only poorer anxiety outcomes but also poorer transdiagnostic outcomes. It could be argued that a fully transdiagnostic approach is likely most beneficial in certain clinical settings. Fully transdiagnostic staging models could be particularly relevant if the goal is early intervention.[28] Fully transdiagnostic models can however suffer from loss of information with regard to certain endophenotypes related to end-stage disorders. Fully transdiagnostic models will also be less easily augmented with stage-specific prognostic or predictive markers as these markers are usually assessed in relation to a specific disorder. Therefore, in the context of multidisciplinary treatment settings aimed at anxiety disorders a half transdiagnostic staging model could be optimal.

Finally, a few other aspects of staging models remain unelucidated. Some authors argue that stage 0 should be used to demarcate at-risk persons,[19] while others argue that the current state of research does not warrant the use of a stage 0.[11] Stage 0 refers to persons without current psychopathology and could prove relevant if these persons have certain susceptibility markers that increase their risks of developing a psychiatric disorder. It is clear that many of such susceptibility markers exist. In addition to the clinical and sociodemographic susceptibility markers that were already mentioned and incorporated into the staging model presented in chapter 5 it is highly likely that genetic factors underly a certain susceptibility for development of anxiety disorders.[29] Due to the presence of susceptibility markers for anxiety disorders the inclusion of stage 0 for at-risk persons seems empirically sound. Whether inclusion of persons in stage 0 should always warrant further evaluation or treatment is however very debatable. From a primary prevention perspective it might be beneficial to apply cost-effective early interventions which could lead to reduced overall burden of mental health in the long term.[28]

Just like the debate revolving stage 0, the operationalization for stage 4 is not fully elucidated. Stage 4 should refer to the patients who exhibited the least favourable clinical course and are subsequently least likely to benefit from prolonged treatments. Stage 4 in the staging model should therefore overlap with the concept of treatment resistance. As was showed in chapter 4, however, the concept of treatment resistance in anxiety disorders is not clearly defined. Possibly, applying the consensus definition for TR-AD that was presented in chapter 4 to stage 4 of the clinical staging model leads to a more clear description of stage 4 in the anxiety disorder staging model.

Ultimately, using clinical staging models has the potential to improve the treatment decision processes in clinical care as different stages could benefit from different treatments. There is evidence for the assumption that interventions in early stages will be both more effective and less harmful than treatments delivered later in the course.[30] These early interventions don't have to be disorder-specific. For instance, many transdiagnostic cognitive processing errors can be treated with cognitive behavioural therapy, thereby reducing the severity and maintenance of symptoms across a range of clinical presentations.[30] Early treatments in prodromal stages of anxiety disorders are likewise effective: cognitive behavioural and psychosocial interventions in high risk groups led to lower anxiety symptoms and lower subsequent incidence of anxiety disorders.[31] Besides the use in

clinical practice, clinical staging may also be useful in research, especially research at the aetiology of anxiety disorders. Possibly, subgroups of anxiety disorders based on the staging model can provide a better basis for research at the aetiology in comparison with using DSM classifications.

Prediction models in anxiety disorders

The second main aim for this thesis was to improve predictions for the clinical course of anxiety disorders by assessing predictive properties of different prediction models. Prediction models are statistical models based on various predictor variables and they provide diagnostic, prognostic or predictive probabilities for certain outcomes.[32] In three of the papers presented in this thesis prediction models were designed. The first model presented is the clinical staging model (chapter 5). This prognostic model is based on a clinician-opinion model for disease progression in anxiety disorders. The second model presented (chapter 6) is a predictive model in which a dimensional measurement instrument was developed based on literature-derived criteria for advanced progression in anxiety disorders. The final prognostic model presented is a data-driven model (chapter 7) for naturalistic course in anxiety disorders. Different statistical metrics are used to evaluate the validity of predictions. The most used metrics are sensitivity and specificity, indicating the proportion of correct predictions among groups that have a certain outcome (sensitivity) versus correct predictions in the group that does not have that outcome (specificity). These can be summarized in either an overall accuracy measure or in an Area Under the Receiving Operator Curve (AUC).[33] To calculate the AUC, all possible predicted probabilities from the prediction model are used as separate thresholds. For each threshold, the predicted probabilities are dichotomized into those above and those below the threshold. Subjects with a predicted probability above the threshold are classified as high risk, while those with predicted probabilities below the threshold are classified as low risk. Using these risk predictions it is possible to calculate sensitivity and specificity for each of these thresholds. The ROC curve is the plot of sensitivity vs. one minus specificity calculated for all possible thresholds. When sensitivity and specificity are high, the AUC is high as well, indicating adequate predictive properties.[34] When predicting binary outcomes, the concordance-statistic (c-statistic) refers to the probability that a randomly selected subject who experienced this outcome will have a higher predicted risk based on the prediction model compared to a randomly selected subject who did not have the outcome.[34] The c-statistic is equal to the AUC. For all prediction models in this thesis, AUCs were calculated to assess accuracy of predictions.

In order to use the clinical staging model as a prediction tool the clinical stages were assessed at baseline according to the adopted model presented in chapter 5. In these analyses, the staging model was coded as a ordinal variable with a range of 0 (stage 0) to 8 (stage 4B). When using this staging model as a prediction tool for 2-year presence of anxiety disorder diagnoses the AUC across all stages was calculated at 0.81 (95% CI: 0.80-0.83). This level of accuracy is substantial. It should be noted, however, that this AUC was calculated when including subjects from preclinical stage 0, 1A and 1B. By including preclinical subjects the accuracy of predictions was inflated at the cost of clinical utility as this comparison is less clinically relevant. A more clinically relevant comparison would be across clinical stages as this could inform clinicians and patients of individualized risks in patients in comparison to other patients. As a post-hoc analysis, the AUC for comparisons across clinical stages 2A to 4B was calculated. This clinical comparison yielded a moderate AUC of 0.64. Therefore, the clinical staging model is moderately able to predict poor outcomes at follow-up. On average, the probabilities are 0.64 that a random anxiety disorder patient who did not remit at two-year follow-up was originally in a higher clinical stage when compared with a random anxiety disorder patient who did remit at two-year follow-up.

The second prediction model assessed is the dimensional measurement instrument based on literature-derived criteria for treatment resistance in anxiety disorders presented in chapter 6. This measurement instrument is based on various clinical characteristics and assessment of these characteristics yields a potential score between 2 and 23, with an observed range in the studied sample of patients with an anxiety disorder who subsequently received treatments from 2 to 16. For each observed score the sensitivity and specificity were calculated for persistence of two-year anxiety disorders and the optimal cut-off value was calculated at 11 or above. This cut-off value was accompanied with a sensitivity of 0.70 and a specificity of 0.57. The AUC for all cut-off values was calculated at 0.66, marginally higher in comparison with the AUC using the staging model. The use of a dimensional measurement instrument for TR-AD was therefore not clearly associated with improved predictive properties for persistence of anxiety disorders at two-year follow-up in comparison with the ordinal staging model.

The final prediction model was the random forests classifier trained in anxiety disorder patients to classify persistence of anxiety disorders at two-year follow-up. This model was trained on a wide array of clinical,

psychological, biological, sociodemographic and lifestyle variables and yielded individual predictions based on decision trees. These predictions were dichotomous: based on the combination of predictor variables a patient was predicted to either be remitted or have a persisting anxiety disorders. Using these individual predictions it was possible to calculate AUCs for the whole model as well as for individual predictor domains. For the whole model the AUC was calculated at 0.67 for persistence of anxiety disorder diagnoses at two-year follow-up. This model thereby provides the highest accuracy of predictions across all three models presented. However, the improvement in predictive properties that was achieved from using a wide array of predictor variables was marginal in comparison with the more straightforward models of clinical staging and the degree of TR-AD measurement tool.

Overall, the accuracy of the predictions in the models presented were consistently moderate. Comparison with earlier prediction models in anxiety disorders are difficult as previous attempts are scarce. One prognostic study into the recurrence of PD in a large sample of remitted PD patients was identified. This study developed a prediction model for recurrence of PD in a sample of 949 remitted PD patients and validated its performance in a sample of 732 remitted PD patients. The prognostic model included eleven predictor variables and the validated predictive properties were adequate with an AUC of 0.73.[35] In a different approach, an European study developed a prediction model for 6-month incidence of GAD and PD in healthy subjects (n=4,905) who visited their general practitioner.[36] The predictive properties for this model were estimated to be substantial with an AUC of 0.78. However, a replication study in a U.S. sample showed poor predictive properties with an AUC of 0.62.[37] A few small machine learning studies focussed on predicting immediate treatment response using neuroimaging data.[38-43] Although some of these studies provided substantially higher AUCs these are not comparable due to the small sample sizes and the cross-sectional design.

There is no guideline for what levels of accuracy in prognostic tests should be deemed adequate. The performance of prognostic tests will depend on different factors. First, the population in which a prediction is tested is highly relevant. If the population consists of a very homogenous group of patients it will become increasingly difficult to adequately predict differences in outcomes among this homogenous group. Alternatively, it is much easier to predict an outcome across a very heterogenous population, as is illustrated by the high predictive performance of the staging model when asymptomatic at-risk subjects were included alongside chronic anxiety disorder patients.

Likewise, it might be easier to predict recurrences in remitted patients as this sample could be more heterogenous in comparison to predicting persistence in current patients. This might explain the somewhat higher AUC in the study into remitted PD patients.[35] Second, the prevalence of outcomes is relevant. For instance, psychiatrists are reasonably well-equipped to predict future suicidal behaviour but their assessments of future suicides are much less precise, partly as this is an outcome with a much lower incidence.[44] This will especially reduce the sensitivity of the prediction as predictions will more likely turn out to be false negative than true positive. Alternatively, in some of the analyses presented in this thesis it appeared more feasible and relevant to predict a broad outcome measure that included all CMDs, so either anxiety disorders, depressive disorders or alcohol dependency as combined outcome measure. By choosing this broad outcome measure the contrast between the two possible outcomes is increased (healthy versus affected). The predictive properties for the ML model were significantly better for this broad outcome measure (AUC increased from 0.67 to 0.70). This increase was largely due to an increase of 7.6 points in sensitivity. The work presented in this thesis thereby highlights the potential benefits of using broad outcome measures in anxiety disorders when studying the longitudinal course. Finally, the amount of time the prognostic test covers is highly relevant. If the amount of time for a prediction increases, the uncertainty in predictions also increases. In all of the models presented, the follow-up periods were extended. This explains the lower predictive performance in the models presented in comparison to the studies assessing immediate treatment response or longitudinal studies with shorter follow-up periods.

Currently diagnostic, prognostic and predictive models are not yet widely implemented. Likely, continued application of machine learning methods to prediction problems in psychiatry will result in development of adequate prediction models. It is shown in other fields of psychiatry that predictions using ML algorithms outperform those using logistic regression analyses.[45] ML prediction models may assist clinical decision making without telling clinicians what to do precisely.

Clinical implications

A number of clinical implications could follow from the results of this thesis. The main implications are diagnostic. First, the results and integration of this thesis warrant applying the staging paradigm in anxiety disorders. As is clear from the studies presented in this thesis, in anxiety disorders, much of the clinical course is defined by the current clinical characteristics. Anxiety

disorder clinical care should therefore at least consist of a combined clinical assessment of previous duration, anxiety severity, presence and severity of functional impairments and presence and severity of somatic and psychiatric comorbidity. After assessing these core clinical characteristics it is possible for the clinician to assess the current clinical stage. The adapted staging model presented in this thesis provides a method for translating these assessments into the appropriate clinical stage.

Second, when assessing anxiety disorder patients, previous treatments and its effects should be assessed. As was clear from the systematic review presented in chapter 4 many authors disregard psychotherapeutic treatments in their assessments of TR-AD. This is not in line with treatment guidelines in which psychotherapeutic treatments play a central role. Currently, stepped-care treatment algorithms for anxiety disorders are based on providing subsequent treatments after treatment failures indicate the need for applying a next step treatment. The definition for TR-AD should therefore be aligned with the current clinical practice. The definition for TR-AD provided in this paper could be used to align diagnostic criteria in this group and provide alignment with treatment guidelines.

Finally, applying prediction models to improve clinical decision-making will become feasible when accuracy of predictions is sufficient. Moreover, prediction models based on predictive markers might be most clinically relevant as they could have direct implications for treatment decisions. For instance, the degree of TR-AD measurement tool could be used to assess likelihood of prospective treatment failures. The findings in this thesis should be seen as encouraging for clinicians interested in personalized medicine.

Future research

Expanding upon the current state of knowledge on prognosis and prediction in anxiety disorders is still needed. In this thesis it was argued that identifying different susceptibility, diagnostic, prognostic and predictive markers across clinical stages of anxiety disorders leads to improvement of evidence-based prognosis and precision psychiatry. Even though stage-specific models are a good method to summarize and visualize the current state of research in anxiety disorders, few studies so far used this stage-specific approach. Future studies could focus on this approach more as it advances evidence-based prognosis and personalized psychiatry in different stages of disease progression. The stage-specific model could still be much more refined by

incorporating research from the wide array of psychiatric research. Many possibly susceptibility, diagnostic, prognostic and predictive markers for anxiety disorder fell outside the scope of the current thesis. For instance, polygenetic risk scores might be linked to increased susceptibility for anxiety disorders, thereby constituting a susceptibility marker.[29] Additionally, certain neuroimaging findings might be useful as diagnostic markers for anxiety disorders. For instance, diagnostic neuroimaging findings in GAD include lower availability of dopamine transporters (DAT) in the striatum, lower number of fronto-cortical GABA-A receptors,[46] larger dorsomedial prefrontal cortical volume in women,[47] and larger amygdala volumes.[48,49] Diagnostic findings in PD include volumetric differences across the basal ganglia and the anterior cingulate cortex.[46] Possible diagnostic markers in SAD include a hyperresponsive emotion network, a diminished cognitive control network, an overactive default-mode network and an active motivational system.[50] Currently, pooled mega-analysis of neuroimaging data is performed worldwide, using data from various sites in the ENIGMA study.[48] Results for anxiety disorders neuroimaging mega-analyses are to be expected shortly. These results could lead to pruning of previous neuroimaging findings and the stage-specific model should be aligned with results from ENIGMA in the future. Other putative diagnostic markers for anxiety disorders that warrant a closer investigation include increased hair cortisol levels in GAD,[51] lower oxytocin and testosterone levels in SAD.[52–54] Behavioural inhibition might be seen as an additional diagnostic marker for at-risk stages of SAD. Behavioural inhibition refers to a particular temperamental trait defined by an inhibited pattern of emotional and behavioural responses to unfamiliar people or unusual situations and incidence of SAD is increased two- to sevenfold in children who show behavioural inhibition in comparison to non-inhibited children.[55,56] Additional possible prognostic markers that were not investigated in this thesis include hypersensitivity to carbon dioxide (CO_2), aberrant levels of tetranectin, creatine kinase MB, ghrelin and lipids in PD.[57–59] Assessing CO_2 reactivity in PD patients after treatment with antidepressants could possibly be a predictive marker and further investigation is warranted as preliminary results suggest that decreased CO_2 reactivity after one week of treatment predicted treatment results after 1 month.[60] Other possible predictive markers include phosphate and BDNF levels for outcomes after CBT in PD,[61,62] as well as altered heart rate variability after treatment with mirtazapine or exposure therapy in PD,[63,64] reversible MAOA gene hypomethylation in PD after CBT,[65] and higher reactivity to fearful faces in the rostral ACC, lesser reactivity in the amygdala and increased activation in the pregenual ACC in PD patients with beneficial treatment effects with

venlafaxine.[66,67] In SAD, right frontal electroencephalography asymmetry might be a predictive markers as it predicted CBT effects without being linked to pre-treatment SAD severity.[68] Furthremore, pharmacogenetic differences might be used as predictive markers as better treatment results with venlafaxine were found in GAD patients with at least one G-allele of the Serotonin receptor 2A.[70] All of these putative predictive markers should be investigated further to assess their merits as expansions to the stage-specific model for anxiety disorders. Novel research modalities might derive further additions to the stage-specific model. Some promising examples include experience sampling methods, [71] gait analysis,[72] actigraphy data,[73] or social media data.[74]

Additionally, in order to fully use the stage-specific model, further investigations in different populations are warranted. For instance, investigating populations of at-risk or subthreshold adolescents might lead to further knowledge on stage 1 markers. Also, studies into treatment resistant samples are needed to further elucidate markers for stage 4 anxiety disorders. From a methodological perspective, the design of the studies presented in this thesis was not adequate to study next-step strategies after applying the prediction models. Future studies could use a prospective longitudinal approach to assess susceptibility and prognostic markers and could use randomized controlled trials to validly assess predictive markers.

In addition, integration of findings from various studies is also much needed. In the second part of this thesis a number of prognostic and predictive markers were integrated into statistical models. The development of such multi-modal models is much needed as it is abundantly clear that no one single factor is sufficiently related to the clinical course in anxiety disorders. However, many iterations are likely needed in order to optimize the statistical layout of these models. Also, frequent updates to these models are warranted as a response to ongoing advances in the scientific field. Clinical utility for prediction models could be increased by performing test reclassification analyses in addition to calculating accuracy and AUCs. In these analyses, the degree to which the use of a prognostic test results in reclassifying a patient into a different risk category is assessed. If high numbers of reclassification occur, the prognostic test adds something to routine clinical care and makes implementation more warranted.[33] Moreover, implementation studies for statistical models should be performed. Also, when studying anxiety disorders the investigated outcomes should include symptoms of other CMDs. The longitudinal analyses in this thesis clearly

showed that the predictive properties for statistical models that included a broad outcome perspective were superior to those that were modelled on a narrow outcome perspective. The aim of evolving statistical prediction models that is mentioned in this thesis is an example of the expanding reliability on advanced computational science. The field of medicine is on the verge of large-scale application of machine learning models.[75] However, some methodological and ethical considerations should be addressed. Currently, only 10.4% of the total models developed in the field of psychiatry are internally validated, and only 4.6% are externally validated.[76] Furthermore, issues regarding patient privacy and confidentiality, informed consent, and patient autonomy were raised.[77] These ethical questions should be continuously asked in order for translation of machine learning models to be implemented effectively and ethically.

Overall, the findings from this thesis point towards the importance of a clear and rigorous diagnostic work-up when assessing a patient with an anxiety disorder. Many different clinical aspects of anxiety disorders are related to the subsequent clinical course. Furthermore, the stage-specific approach in identifying diagnostic, prognostic and predictive risk factors is a valuable addition to the diagnostic framework. And although the predictive properties should still be refined, this thesis provides different easy to use diagnostic, prognostic and predictive models for clinicians. This focus on diagnostic, prognostic and predictive aspects of anxiety disorders is an important step toward precision psychiatry and could lead to improvements in the treatment decision process, leading to providing patients with the right treatment at the right time.

References

1. Penninx BWJH, Nolen WA, Lamers F, et al. Two-year course of depressive and anxiety disorders: Results from the Netherlands Study of Depression and Anxiety (NESDA). *J Affect Disord*. 2011;133(1-2):76-85. doi:10.1016/j.jad.2011.03.027
2. Amur S, Lavange L, Zineh I, Buckman-Garner S, Woodcock J. Biomarker qualification: Toward a multiple stakeholder framework for biomarker development, regulatory acceptance, and utilization. *Clin Pharmacol Ther*. 2015;98(1):34-46. doi:10.1002/cpt.136
3. Cosci F, Mansueto G. Biological and clinical markers in panic disorder. *Psychiatry Investig*. 2019;16(1):27-36. doi:10.30773/pi.2018.07.26
4. Stein MB, Sareen J. Generalized Anxiety Disorder. *N Engl J Med*. 2015;373(21):2059-2068. doi:10.1056/NEJMcp1502514
5. Moreno-Peral P, Conejo-Cerón S, Motrico E, et al. Risk factors for the onset of panic and generalised anxiety disorders in the general adult population: A systematic review of cohort studies. *J Affect Disord*. 2014;168:337-348. doi:10.1016/j.jad.2014.06.021
6. Hyland P, Shevlin M, Elklit A, Christoffersen M, Murphy J. Social, familial and psychological risk factors for mood and anxiety disorders in childhood and early adulthood: a birth cohort study using the Danish Registry System. *Soc Psychiatry Psychiatr Epidemiol*. 2016;51(3):331-338. doi:10.1007/s00127-016-1171-1
7. Fass R, Achem SR. Noncardiac chest pain: Epidemiology, natural course and pathogenesis. *J Neurogastroenterol Motil*. 2011;17(2):110-123. doi:10.5056/jnm.2011.17.2.110
8. Kisely SR, Campbell LA, Yelland MJ, Paydar A. Psychological interventions for symptomatic management of non-specific chest pain in patients with normal coronary anatomy. *Cochrane Database Syst Rev*. 2015;2015(6). doi:10.1002/14651858.CD004101.pub5
9. Batelaan NM, Smit JH, Cuijpers P, van Marwijk HWJ, Terluin B, van Balkom AJLM. Prevention of anxiety disorders in primary care: a feasibility study. *BMC Psychiatry*. 2012;12(1):206. doi:10.1186/1471-244X-12-206
10. Lau EX, Rapee RM. Prevention of anxiety disorders. *Curr Psychiatry Rep*. 2011;13(4):258-266. doi:10.1007/s11920-011-0199-x
11. Cosci F, Fava GA. Staging of mental disorders: systematic review. *Psychother Psychosom*. 2013;82(1):20-34. doi:10.1159/000342243
12. Kessler RC, Chiu WT, Demler O, Walters EE, Merikangas KR. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(June 2005):617-627. doi:10.1001/archpsyc.62.6.617

13. Schoevers RA, Deeg DJH, van Tilburg W, Beekman ATF. Depression and Generalized Anxiety Disorder: Co-Occurrence and Longitudinal Patterns in Elderly Patients. *Am J Geriatr Psychiatry*. 2005;13(1):31-39. doi:10.1097/00019442-200501000-00006
14. Ojala M, Garriga GC. Permutation Tests for Studying Classifier Performance. *J Mach Learn Res*. 2010;11:1833-1863.
15. Insel T, Cuthbert B, Garvey M, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry*. 2010;167(July):748-751. doi:10.1176/appi.ajp.2010.09091379
16. Fava GA, Kellner R. Staging: a neglected dimension in psychiatric classification. *Acta Psychiatr Scand*. 1993;87(4):225-230. http://www.ncbi.nlm.nih.gov/pubmed/8488741. Accessed November 21, 2014.
17. Langenbucher JW, Chung T. Onset and staging of DSM-IV alcohol dependence using mean age and survival-hazard methods. *J Abnorm Psychol*. 1995;104(2):346-354.
18. McGorry PD, Hickie IB, Yung AR, Pantelis C, Jackson HJ. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. *Aust N Z J Psychiatry*. 2006;40(8):616-622. doi:10.1111/j.1440-1614.2006.01860.x
19. Hickie IB, Scott EM, Hermens DF, et al. Applying clinical staging to young people who present for mental health care. *Early Interv Psychiatry*. 2013;7(1):31-43. doi:10.1111/j.1751-7893.2012.00366.x
20. Johnson EM, Coles ME. Failure and delay in treatment-seeking across anxiety disorders. *Community Ment Health J*. 2013;49(6):668-674.
21. Beesdo-Baum K, Knappe S, Fehm L, et al. The natural course of social anxiety disorder among adolescents and young adults. *Acta Psychiatr Scand*. 2012;126(6):411-425. doi:10.1111/j.1600-0447.2012.01886.x
22. Ginsburg GS, Becker EM, Keeton CP, et al. Naturalistic follow-up of youths treated for pediatric anxiety disorders. *JAMA psychiatry*. 2014;71(3):310-318. doi:10.1001/jamapsychiatry.2013.4186
23. Lamers F, van Oppen P, Comijs HC, et al. Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). *J Clin Psychiatry*. 2011;72(3):341-348. doi:10.4088/JCP.10m06176blu
24. Hovenkamp-Hermelink JHM, Riese H, Van Der Veen DC, Batelaan NM, Penninx BWJH, Schoevers RA. Low stability of diagnostic classifications of anxiety disorders over time: A six-year follow-up of the NESDA study. *J Affect Disord*. 2016;190:310-315. doi:10.1016/j.jad.2015.10.035

25. Otowa T, Hek K, Lee M, et al. Meta-analysis of genome-wide association studies of anxiety disorders. *Mol Psychiatry*. 2016;21(10):1391-1399. doi:10.1038/mp.2015.197
26. Craske MG, Stein MB, Eley TC, et al. Anxiety disorders. *Nat Rev Dis Prim*. 2017;3:17-24. doi:10.1038/nrdp.2017.24
27. Scholten WD, Batelaan NM, Penninx BWJH, et al. Diagnostic instability of recurrence and the impact on recurrence rates in depressive and anxiety disorders. *J Affect Disord*. 2016;195:185-190. doi:10.1016/j.jad.2016.02.025
28. McGorry PD, Ratheesh A, O'Donoghue B. Early Intervention—an Implementation Challenge for 21st Century Mental Health Care. *JAMA Psychiatry*. 2018;75(6):545-546. doi:10.1001/jamapsychiatry.2018.0621
29. García-Gutiérrez MS, Navarrete F, Sala F, Gasparyan A, Austrich-Olivares A, Manzanares J. Biomarkers in Psychiatry: Concept, Definition, Types and Relevance to the Clinical Reality. *Front Psychiatry*. 2020;11(May):1-14. doi:10.3389/fpsyt.2020.00432
30. Scott J, Leboyer M, Hickie I, et al. Clinical staging in psychiatry: a cross-cutting model of diagnosis with heuristic and practical value. *Br J Psychiatry*. 2013;202(4):243-245. doi:10.1192/bjp.bp.112.110858
31. Brenner R, Madhusoodanan S, Puttichanda S, Chandra P. Primary prevention in psychiatry--adult populations. *Ann Clin Psychiatry*. 2010;22(4):239-248.
32. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. (Gail M, Krickeberg K, Samet J, Tsiatis A, Wong W, eds.). New York, NY, US: Springer; 2009. doi:10.1007/978-1-4757-3294-8_4
33. Rector TS, Taylor BC, Wilt TJ. *Systematic Review of Prognostic Tests.* (Chang SM, Matchar DB, Smetana GW, Umscheid CA, Gray R, Torchia MM, eds.). Rockville, MD: Agency for Healthcare Research and Quality; 2012.
34. Austin PC, Steyerberg EW. Interpreting the concordance statistic of a logistic regression model: Relation to the variance and odds ratio of a continuous explanatory variable. *BMC Med Res Methodol*. 2012;12:1-8. doi:10.1186/1471-2288-12-82
35. Liu Y, Sareen J, Bolton J, Wang J. Development and validation of a risk-prediction algorithm for the recurrence of panic disorder. *Depress Anxiety*. 2015;32(5):341-348. doi:10.1002/da.22359
36. King M, Bottomley C, Bellón-Saameño JA, et al. An international risk prediction algorithm for the onset of generalized anxiety and panic syndromes in general practice attendees: PredictA. *Psychol Med*. 2011;41(8):1625-1639. doi:10.1017/S0033291710002400

37. Liu Y, Sareen J, Bolton JM, Wang JL. Development and validation of a risk prediction algorithm for the recurrence of suicidal ideation among general population with low mood. *J Affect Disord*. 2016;193:11-17. doi:10.1016/j.jad.2015.12.072
38. Månsson KN, Frick A, Boraxbekk C-JJ, et al. Predicting long-term outcome of Internet-delivered cognitive behavior therapy for social anxiety disorder using fMRI and support vector machine learning. *Transl Psychiatry*. 2015;5:e530. doi:10.1038/tp.2015.22
39. Ball TM, Stein MB, Ramsawh HJ, Campbell-Sills L, Paulus MP. Single-subject anxiety treatment outcome prediction using functional neuroimaging. *Neuropsychopharmacology*. 2014;39(5):1254-1261. doi:10.1038/npp.2013.328
40. Doehrmann O, Ghosh SS, Polli FE, et al. Predicting treatment response in social anxiety disorder from functional magnetic resonance imaging. *JAMA Psychiatry*. 2013;70(1):87-97. doi:10.1001/2013.jamapsychiatry.5
41. Hahn T, Kircher T, Straube B, et al. Predicting treatment response to cognitive behavioral therapy in panic disorder with agoraphobia by integrating local neural information. *JAMA Psychiatry*. 2015;72(1):68-74. doi:10.1001/jamapsychiatry.2014.1741
42. Pantazatos SP, Talati A, Schneier FR, Hirsch J. Reduced anterior temporal and hippocampal functional connectivity during face processing discriminates individuals with social anxiety disorder from healthy controls and panic disorder, and increases following treatment. *Neuropsychopharmacology*. 2014;39(2):425-434. doi:10.1038/npp.2013.211
43. Whitfield-Gabrieli S, Ghosh SS, Nieto-Castanon A, et al. Brain connectomics predict response to treatment in social anxiety disorder. *Mol Psychiatry*. 2016;21(5):680-685. doi:10.1038/mp.2015.109
44. Randall JR, Sareen J, Chateau D, Bolton JM. Predicting Future Suicide: Clinician Opinion versus a Standardized Assessment Tool. *Suicide Life-Threatening Behav*. 2019;49(4):941-951. doi:10.1111/sltb.12481
45. Kessler RC, van Loo HM, Wardenaar KJ, et al. Testing a machine-learning algorithm to predict the persistence and severity of major depressive disorder from baseline self-reports. *Mol Psychiatry*. 2016;21(10):1366-1371. doi:10.1038/mp.2015.198
46. Bandelow B, Baldwin D, Abelli M, et al. Biological markers for anxiety disorders, OCD and PTSD: A consensus statement. Part II: Neurochemistry, neurophysiology and neurocognition. *World J Biol Psychiatry*. 2017;18(3):162-214. doi:10.1080/15622975.2016.1190867
47. Schienle A, Ebner F, Schäfer A. Localized gray matter volume abnormalities in generalized anxiety disorder. *Eur Arch Psychiatry Clin Neurosci*. 2011;261(4):303-307. doi:10.1007/s00406-010-0147-5

48. Makovac E, Meeten F, Watson DR, Garfinkel SN, Critchley HD, Ottaviani C. Neurostructural abnormalities associated with axes of emotion dysregulation in generalized anxiety. *NeuroImage Clin*. 2016;10(December):172-181. doi:10.1016/j.nicl.2015.11.022
49. Etkin A, Prater KE, Schatzberg AF, Menon V, Greicius MD. Disrupted amygdalar subregion functional connectivity and evidence of a compensatory network in generalized anxiety disorder. *Arch Gen Psychiatry*. 2009;66(12):1361-1372. doi:10.1001/archgenpsychiatry.2009.104
50. Cremers HR, Roelofs K. Social anxiety disorder: a critical overview of neurocognitive research. *Wiley Interdiscip Rev Cogn Sci*. 2016;7(4):218-232. doi:10.1002/wcs.1390
51. Steudte S, Stalder T, Dettenborn L, et al. Decreased hair cortisol concentrations in generalised anxiety disorder. *Psychiatry Res*. 2011;186(2-3):310-314. doi:10.1016/j.psychres.2010.09.002
52. Hoge EA, Lawson EA, Metcalf CA, et al. Plasma oxytocin immunoreactive products and response to trust in patients with social anxiety disorder. *Depress Anxiety*. 2012;29(11):924-930. doi:10.1002/da.21973
53. Hoge EA, Pollack MH, Kaufman RE, Zak PJ, Simon NM. Oxytocin levels in social anxiety disorder. *CNS Neurosci Ther*. 2008;14(3):165-170. doi:10.1111/j.1755-5949.2008.00051.x
54. Giltay EJ, Enter D, Zitman FG, et al. Salivary testosterone: Associations with depression, anxiety disorders, and antidepressant use in a large cohort study. *J Psychosom Res*. 2012;72(3):205-213. doi:10.1016/j.jpsychores.2011.11.014
55. Hirshfeld-Becker DR, Biederman J, Henin A, et al. Behavioral inhibition in preschool children at risk is a specific predictor of middle childhood social anxiety: A five-year follow-up. *J Dev Behav Pediatr*. 2007;28(3):225-233. doi:10.1097/01.DBP.0000268559.34463.d0
56. Clauss JA, Blackford JU. Behavioral inhibition and risk for developing social anxiety disorder: A meta-analytic study. *J Am Acad Child Adolesc Psychiatry*. 2012;51(10):1066-1075.e1. doi:10.1016/j.jaac.2012.08.002
57. Perna G, Barbini B, Cocchi S, Bertani A, Gasperini M. 35% CO2 challenge in panic and mood disorders. *J Affect Disord*. 1995;33(3):189-194. doi:10.1016/0165-0327(94)00088-Q
58. Gottschalk MG, Cooper JD, Chan MK, Bot M, Penninx BWJH, Bahn S. Serum biomarkers predictive of depressive episodes in panic disorder. *J Psychiatr Res*. 2016;73:53-62. doi:10.1016/j.jpsychires.2015.11.012
59. Gul IG, Cumurcu BE, Karlidag R, Turkoz Y. Ghrelin and lipid levels in panic disorder before and after treatment and their relationship with agoraphobia. *Psychiatr Danub*. 2015;27(3):250-258.

60. Ghaziuddin N, Dumas S, Hodges E. Use of continuation or maintenance electroconvulsive therapy in adolescents with severe treatment-resistant depression. *J ECT*. 2011;27(2):168-174. doi:10.1097/YCT.0b013e3181f665e4
61. Kobayashi K, Shimizu E, Hashimoto K, et al. Serum brain-derived neurotrophic factor (BDNF) levels in patients with panic disorder: As a biological predictor of response to group cognitive behavioral therapy. *Prog Neuro-Psychopharmacology Biol Psychiatry*. 2005;29(5):658-663. doi:10.1016/j.pnpbp.2005.04.010
62. Beria P, Viana ACW, Behenck A, Heldt E, Manfro GG, Dreher CB. Respiratory subtype of panic disorder: Can serum phosphate levels be a possible outcome to group cognitive-behavior therapy? *J Affect Disord*. 2018;235(November 2017):474-479. doi:10.1016/j.jad.2018.04.002
63. Slaap BR, Boshuisen ML, Van Roon AM, Den Boer JA. Heart rate variability as predictor of nonresponse to mirtazapine in panic disorder: A preliminary study. *Int Clin Psychopharmacol*. 2002;17(2):69-74.
64. Wendt J, Hamm AO, Pane-Farre CA, et al. Pretreatment Cardiac Vagal Tone Predicts Dropout from and Residual Symptoms after Exposure Therapy in Patients with Panic Disorder and Agoraphobia. *Psychother Psychosom*. 2018. doi:10.1159/000487599
65. Ziegler C, Richter J, Mahr M, et al. MAOA gene hypomethylation in panic disorder - Reversibility of an epigenetic risk pattern by psychotherapy. *Transl Psychiatry*. 2016;6. doi:10.1038/tp.2016.41
66. Whalen PJ, Johnstone T, Somerville LH, et al. A Functional Magnetic Resonance Imaging Predictor of Treatment Response to Venlafaxine in Generalized Anxiety Disorder. *Biol Psychiatry*. 2008;63(9):858-863. doi:10.1016/j.biopsych.2007.08.019.A
67. Nitschke JB, Sarinopoulos I, Oathes DJ, et al. Anticipatory Activation in the Amygdala and Anterior Cingulate in Generalized Anxiety Disorder and Prediction of Treatment Response. *Am J Psychiatry*. 2009;166(3):302-320.
68. Shin LM, Davis F, VanElzakker MB, Dahlgren MK, Dubois SJ. Neuroimaging predictors of treatment response in anxiety disorders. *Biol Mood Anxiety Disord*. 2013;3(1):15. doi:10.1186/2045-5380-3-15
69. Santos VA, Carvalho DD, Van Ameringen M, Nardi AE, Freire RC. Neuroimaging findings as predictors of treatment outcome of psychotherapy in anxiety disorders. *Prog Neuro-Psychopharmacology Biol Psychiatry*. 2019;91(April 2018):60-71. doi:10.1016/j.pnpbp.2018.04.001
70. Lohoff FW, Aquino TD, Narasimhan S, Multani PK, Etemad B, Rickels K. Serotonin receptor 2A (HTR2A) gene polymorphism predicts treatment response to venlafaxine XR in generalized anxiety disorder. *Pharmacogenomics J*. 2013;13(1):21-26. doi:10.1038/tpj.2011.47

71. Kubben P, Dumontier M, Dekker A. *Fundamentals of Clinical Data Science*. Cham, Switzerland: Springer Open; 2019.
72. Zhao N, Zhang Z, Wang Y, et al. See your mental state from your walk: Recognizing anxiety and depression through Kinect-recorded gait data. *PLoS One*. 2019;14(5):1-13. doi:10.1371/journal.pone.0216591
73. Merikangas KR, Swendsen J, Hickie IB, et al. Real-time Mobile Monitoring of the Dynamic Associations Among Motor Activity, Energy, Mood, and Sleep in Adults With Bipolar Disorder. *JAMA Psychiatry*. 2019;Feb:190-198. doi:10.1001/jamapsychiatry.2018.3546
74. Reece AG, Danforth CM. Instagram photos reveal predictive markers of depression. *EPJ Data Sci*. 2017;6(1). doi:10.1140/epjds/s13688-017-0110-z
75. Shah P, Kendall F, Khozin S, et al. Artificial intelligence and machine learning in clinical development: a translational perspective. *npj Digit Med*. 2019;2(1). doi:10.1038/s41746-019-0148-3
76. Salazar de Pablo G, Studerus E, Vaquerizo-Serrano J, et al. Implementing Precision Psychiatry: A Systematic Review of Individualized Prediction Models for Clinical Practice. *Schizophr Bull*. 2020:1-14. doi:10.1093/schbul/sbaa120
77. Rigby MJ. Ethical dimensions of using artificial intelligence in health care. *AMA J Ethics*. 2019;21(2):121-124. doi:10.1001/amajethics.2019.121

Addendum

Samenvatting

Het doel van **hoofdstuk 2** was om een screeningsprogramma voor angststoornissen te implementeren bij patiënten die met acute pijn op de borst de eerste harthulp bezochten. 'Niet-cardiale pijn op de borst' wordt gediagnosticeerd bij patiënten met acute pijn op de borst nadat een cardiale oorzaak van de pijn is uitgesloten. Meer dan de helft van alle patiënten met pijn op de borst krijgt de diagnose niet-cardiale pijn op de borst. Onderzoek wijst uit dat een aanzienlijk deel (12-41%) van deze patiënten een paniekstoornis heeft. In de praktijk wordt tot nu toe niet gescreend op paniekstoornissen en andere psychiatrische stoornissen bij patiënten met niet-cardiale pijn op de borst. Vanwege de hoge prevalentiecijfers van deze stoornissen is screening belangrijk. Het doel van deze studie was om te screenen op paniekstoornissen en andere psychiatrische stoornissen bij patiënten met niet-cardiale pijn op de borst. Het onderzoek werd tussen 2012 en 2013 uitgevoerd op de eerste harthulp van het VU Medisch Centrum. Een groep van 252 volwassen patiënten met niet-cardiale pijn op de borst kwam in aanmerking voor screening naar psychiatrische stoornissen. Het screeningsprogramma bestond uit twee fasen: de eerste fase werd uitgevoerd op de eerste harthulp door getrainde cardiologie verpleegkundigen en bestond uit het afnemen van de *Hospital Anxiety and Depression Scale* (HADS) nadat de diagnose niet-cardiale pijn op de borst was gesteld. Alle patiënten die angstige of depressieve symptomen vertoonden boven een afkapwaarde op de HADS kwamen in aanmerking voor de tweede fase van de screening. Deze tweede fase werd ingepland op een later tijdstip en vond plaats op de afdeling psychiatrie in hetzelfde ziekenhuis. De tweede screeningsfase bestond uit een gestructureerd interview waarbij verschillende psychiatrische diagnoses werden onderzocht met behulp van het *Composite International Diagnostic Interview* (CIDI).

Helaas werd de eerste fase van screening bij slechts 60 van de 252 (23,8%) patiënten met niet-cardiale pijn op de borst uitgevoerd. Dit lage aantal was grotendeels te wijten aan een lage bereidheid van verpleegkundigen om het screeningsprogramma op te starten. Zij gaven aan dat andere taken een hogere prioriteit hadden dan het screeningsprogramma. Van de 60 patiënten die screening kregen aangeboden op de eerste harthulp waren 51 patiënten (85,0%) bereid hieraan deel te nemen. Bijna de helft van de gescreende patiënten met niet-cardiale pijn op de borst (24/51; 47,1%) scoorde boven de afkapwaarde op de HADS en kwam in aanmerking voor de tweede fase van screening. In de aanloop naar de tweede fase van screening

zagen 12 patiënten (50%) af van verdere deelname. Van de overige 12 patiënten werden 8 patiënten gediagnosticeerd met een psychiatrische stoornis, waaronder twee met de diagnose paniekstoornis. Gezien deze lage aantallen werd verdere implementatie van het screeningsprogramma niet haalbaar geacht. De belangrijkste barrière voor implementatie was een onmogelijkheid van de verpleegkundigen om het uitvoeren van de screening te combineren met de andere taken in hun takenpakket.

Het doel van **hoofdstuk 3** was om de impact van angst- en depressieve stoornissen en chronische somatische ziektes op beperkingen en beroepsmatig functioneren te beoordelen. Patiënten met angst- en depressieve stoornissen hebben vaak functionele beperkingen en beperkingen in het arbeidsmatige functioneren. Hetzelfde geldt voor patiënten met chronische somatische ziektes, zoals lage rugpijn, migraine, diabetes en obesitas. Bovendien komen angst- en depressieve stoornissen vaak samen voor met chronische somatische ziektes. De comorbiditeit tussen deze psychiatrische aandoeningen met somatische aandoeningen vermindert het algehele functioneren nog verder en er wordt aangenomen dat het de behandelresultaten voor somatische ziekten vermindert. Verder hebben chronische somatische ziektes waarschijnlijk een negatieve invloed op het beloop van angst- en depressieve stoornissen. Het doel van deze studie was om de effecten van angst- en depressieve stoornissen, chronische somatische ziektes en comorbiditeit hiertussen op functionele beperkingen bij volwassen patiënten te onderzoeken. We gebruikten de gegevens van de Nederlandse Studie naar Depressie en Angst (NESDA) om patiënten te includeren voor dit onderzoek. De steekproef omvatte 2371 personen, bestaande uit patienten met angst- en depressieve stoornissen en controles. De onderzochte angststoornissen waren de gegeneraliseerde angststoornis, de sociale angststoornis, agorafobie en de paniekstoornis. De depressieve stoornissen omvatten de depressieve stoornis (*major depressive disorder*) en dysthymie. De aanwezigheid van 30 verschillende chronische somatische ziektes werd onderzocht. Somatische ziekten werden alleen meegeteld wanneer patiënten hiervoor een behandeling of medicatie nodig hadden. In de steekproef hadden patiënten met angst- en depressieve stoornissen vaker een chronische somatische ziekte. Uit gecorrigeerde logistische regressieanalyses werd een *odds-ratio* (OR) van 1,34 (95% betrouwbaarheidsinterval (BI) = 1,09-1,64) afgeleid voor patiënten met angst- en depressieve stoornissen voor het hebben van een van de chronische somatische ziektes. De verschillende chronische somatische ziektes werden onderverdeeld in zeven categorieën: respiratoire, cardio-

metabole, musculoskeletale, gastro-intestinale, neurologische, endocriene en kankers. Van deze categorieën bleek de categorie gastro-intestinale aandoeningen de enige die significant vaker aanwezig was bij patiënten met angst- en depressieve stoornissen (OR=3.29, 2.15-5.05).

Uit de beschrijvende statistiek werd duidelijk dat angst- en depressieve stoornissen samenhingen met een lagere functionele status: de totale gestandaardiseerde score voor beperkingen, gemeten met het *World Health Organization Disability Assessment Schedule* (WHO-DAS II) was 29,0 ± 16,4 terwijl de totale score voor beperkingen in de controlegroep 7,8 ± 9,3 was. Multivariate lineaire regressieanalyses toonden aan dat angst- en depressieve stoornissen gerelateerd waren aan de hoogste niveaus van beperkingen (β = 20,1). Chronische somatische ziektes waren in mindere mate ook geassocieerd met beperkingen (β = 3,88). Bovendien was er een interactie-effect aanwezig tussen angst- en depressieve stoornissen en chronische somatische ziektes op de mate van beperkingen (β = 4,06). Uit dit interactiemodel bleek dat de effecten van chronische somatische ziektes op beperkingen vooral aanwezig waren bij patiënten die ook angst- en depressieve stoornissen hebben. Er bleek sprake van synergistische effectmodulatie, waarbij het effect van comorbide chronische somatische ziektes en angst- en depressieve stoornissen op scores voor beperkingen groter is dan de som van de afzonderlijke effecten voor angst- en depressieve stoornissen en chronische somatische ziektes.

Analyses van arbeidsbeperkingen werden uitgevoerd in een subset van 1462 respondenten die aangaven te werken. Uit beschrijvende statistiek bleek dat de arbeidsbeperkingen groter waren bij patiënten met angst- en depressieve stoornissen: 67,6% van de patiënten had op zijn minst enig ziekteverzuim, terwijl slechts 32,7% van de controles ziekteverzuim had (X^2 = 163,9, p<0,001). Evenzo hadden patiënten met angst- en depressieve stoornissen in vergelijking met controles meer kans op verminderde werkprestaties: 60,5% van de patiënten had enige vorm van verminderde werkprestatie terwijl slechts 32,5% van de controles verminderde werkprestaties had (X^2 = 132,2, p<0,001). Er werden multinomiale regressieanalyses uitgevoerd om interactie effecten te kunnen beoordelen. Hierbij werden de metingen ten aanzien van ziekteverzuim en verminderde werkprestaties gecategoriseerd en werden vergelijkingen gemaakt met de 'gezonde' categorieën van geen verzuim en geen verminderde werkprestaties. Het leek erop dat chronische somatische ziektes alleen geassocieerd waren met de meest ernstige uitkomsten van werkbeperkingen: langdurig verzuim OR= 1,42 (95% BI

1,07-1,88) en ernstig verminderde werkprestaties OR= 1,42 (95% BI 1,08-1,87). Angst- en depressieve stoornissen waren geassocieerd met alle uitkomsten van werkbeperkingen: kort verzuim OR= 2,88 (95% BI 2,16-3,84), langdurig verzuim OR= 6,64 (95% BI 4,69-9,40), licht verminderde werkprestaties OR= 1,83 (95% BI 1,38- 2,43) en ernstig verminderde werkprestaties OR= 7,51 (95% BI 5,11-11,1). Deze regressiecoëfficiënten werden gecorrigeerd voor sociodemografische gegevens en laten een duidelijk verschil zien tussen angst- en depressieve stoornissen en chronische somatische ziektes op de uitkomsten van arbeidsbeperkingen, waarbij de effecten van angst- en depressieve stoornissen veel groter waren. Daarnaast werden modellen gemaakt met behulp van vier categorieën voor combinaties van blootstelling aan angst- en depressieve stoornissen en chronische somatische ziektes: controles (geen angst- en depressieve stoornissen, geen chronische somatische ziektes), uitsluitend fysiek (chronische somatische ziektes zonder angst- en depressieve stoornissen), uitsluitend mentaal (angst- en depressieve stoornissen zonder chronische somatische ziektes) en comorbiditeit (angst- en depressieve stoornissen en chronische somatische ziektes). Er was geen relatie tussen chronische somatische ziekte en werkverzuim, maar wel met verminderde werkprestaties. De groep met angst- en depressieve stoornissen bleek te maken te hebben met hogere maten van werkverzuim en verminderde werkprestaties. Ten slotte bleek de groep met comorbiditeit geassocieerd met de slechtste uitkomsten voor alle maten van arbeidsbeperkingen.

Deze bevindingen benadrukken het belang van het herkennen en behandelen van psychiatrische comorbiditeit bij patiënten met chronische somatische ziektes en eveneens van het herkennen en behandelen van somatische comorbiditeit bij patiënten met angst- en depressieve stoornissen om langdurige beperkingen en verminderd beroepsmatig functioneren bij deze patiënten tegen te gaan.

Het doel van **hoofdstuk 4** was om verschillende definities en diagnostische criteria voor therapieresistentie bij angststoornissen (TR-AD) in kaart te brengen. Op dit moment zijn er geen eenduidige criteria voor TR-AD. Het is belangrijk dat er duidelijke criteria komen, omdat de behandeling voor veel patiënten met een angststoornis een beperkt resultaat heeft. In de literatuur worden verschillende termen gebruikt voor het optreden van een suboptimaal behandeleffect. Voorbeelden hiervan zijn 'therapieresistentie', 'symptomatisch blijven', 'refractair' en 'non-respons'. Deze termen worden min of meer inwisselbaar gebruikt. Hoofdstuk

4 beschrijft een systematisch review waarin deze termen in kaart werden gebracht. Alle publicaties die een duidelijke beschrijving van het fenomeen TR-AD bevatten werden meegenomen in het review. De verzamelde beschrijvingen werden systematisch beschreven met betrekking tot verschillende aspecten waar deze uit bestonden. Het doel was om alle verschillende criteria en aspecten van definities van TR-AD te verzamelen en weer te geven. Een secundair doel was om deze criteria te integreren tot een nieuwe en eenduidige consensusdefinitie voor TR-AD.

De zoekstrategie leverde 13.042 unieke records op. Deze publicaties zijn eerst onafhankelijk op titel en abstract beoordeeld door twee onderzoekers. Dit resulteerde in 388 records die in aanmerking kwamen voor full-text screening en daarvan werden 62 studies opgenomen in de datasynthese. De geïncludeerde onderzoeken gaven allemaal een specifieke definitie voor TR-AD, of ze gaven inclusiecriteria voor patiënten die werden beschreven als TR-AD-patiënten. De selectie van studies bestond uit reviews, richtlijnen, boekhoofdstukken, trials en cohortstudies. De bestudeerde angststoornissen waren de paniekstoornis (n=33), de gegeneraliseerde angststoornis (n=34), de sociale angststoornis (n=21), de specifieke fobie (n=5) en angststoornissen in het algemeen (n=5). Voordat een patiënt aan de criteria voor TR-AD voldoet, zijn er volgens de meeste onderzoeken (85,5%) een minimaal aantal behandelingen zonder resultaat nodig. Het minimale aantal behandelingen met onvoldoende resultaat varieerde van één tot vijf. Verdere criteria waren onder meer de voorwaarde dat er farmacologische behandelingen zonder resultaat moesten zijn (93,5%), de voorwaarde dat er psychotherapieën met onvoldoende resultaat moeten zijn (29,0%), de voorwaarde dat deze behandelingen een minimale duur moesten omvatten (54,8%), de voorwaarde dat een behandeling pas mocht worden beschouwd als één met onvoldoende resultaat wanneer er niet werd voldaan aan een specifiek responscriterium (41,9%), bijv. een vermindering van de ernst van de angstklachten van meer dan 50%. Bovendien werd in bijna de helft (46,8%) van de onderzoeken een drempel voor de ernst van angstklachten gebruikt vóór behandeling. Ten slotte werden enkele voorwaarden in een kleine minderheid (minder dan 10%) gebruikt in TR-AD-definities: minimale duur van de angststoornis, de aanwezigheid van functionele beperkingen en de aanwezigheid van psychiatrische comorbiditeit. Uit kwaliteitsanalyses bleek dat de kwaliteit van het onderzoek geen invloed had op de door de auteurs gebruikte criteria voor TR-AD: dezelfde criteria werden gerapporteerd bij onderzoeken van lage en hoge kwaliteit.

Ten slotte werden de gevonden criteria uit deze systematische review geïntegreerd in een nieuwe consensusdefinitie voor TR-AD. Hiervoor werden de meest gebruikte criteria en aspecten voor TR-AD geselecteerd. Volgens deze consensusdefinitie is TR-AD aanwezig als patiënten ten minste één farmacologische behandeling met een eerstelijns antidepressivum (SSRI/SNRI) hadden. Daarnaast moeten patiënten ook minimaal één eerstelijns protocollaire psychotherapeutische behandeling (CGT) gericht op de angststoornis hebben gehad. Deze behandelingen tellen alleen mee wanneer de behandeling er niet in slaagde om een angstreductie van minimaal 50% te bereiken na een behandelingsduur van minimaal 8 weken. Ten slotte moet de ernst van de symptomen boven een gespecificeerde drempel liggen om aan de voorwaarden voor TR-AD te kunnen voldoen. De afkapwaarden waren per angststoornis verschillend en voor de consensusdefinitie kozen wij de meest gebruikte afkapwaarden. Het consequent gebruiken van deze nieuwe consensusdefinitie in onderzoeken naar TR-AD zal de homogeniteit van de bestudeerde populaties vergroten en de generaliseerbaarheid van bevindingen vergroten, waardoor er effectiever onderzoek kan worden gedaan naar vervolgstappen bij TR-AD.

Het doel van **hoofdstuk 5** was om een bestaand generiek stadiëringsmodel voor psychiatrische stoornissen te vertalen voor gebruik bij patiënten met angststoornissen. De belangrijkste doelen waren het beoordelen van de construct validiteit en de voorspellende validiteit van deze aangepaste versie van dit stadiëringsmodel. In het paradigma van klinische stadiëring kunnen verschillende stadia worden onderscheiden die de toenemende mate van ziekteprogressie weerspiegelen. Elk hoger stadium gaat gepaard met een minder gunstig verder ziekteverloop. Theoretisch zouden de verschillende stadia in een stadiëringsmodel ook verschillende onderliggende pathofysiologische processen moeten weerspiegelen. In de psychiatrie zijn de onderliggende pathofysiologische kenmerken multifactorieel en hierdoor minder geschikt voor deze benadering. Om die reden zijn de huidige stadiëringsmodellen in de psychiatrie voornamelijk gebaseerd op klinische kenmerken.

Voor deze studie werd een stadiëringsmodel van een hoog aangeschreven Australische onderzoeksgroep aangepast voor gebruik bij angststoornissen. Dit leverde een stadiëringsmodel op met stadia die variëren van stadium 0 (asymptomatisch, hoog risico) tot stadium 4B (chronische symptomen met comorbiditeit). Op de baselinemeting van NESDA werden proefpersonen geïncludeerd in dit onderzoek. Alle 1305

NESDA-proefpersonen met een angststoornis en 1115 proefpersonen zonder huidige angststoornis of depressieve stoornis maar met aanwezigheid van risicofactoren voor het ontwikkelen van een angststoornis werden geïncludeerd. Alle 2420 geïncludeerde proefpersonen werden toegewezen aan een klinisch stadium. De toewijzing aan stadia werd gebaseerd op *life chart* interviews die betrekking hadden op de duur van de angststoornis. Verder werd de ernst van de symptomen gemeten met vragenlijsten en werd de aanwezigheid van psychiatrische comorbiditeit onderzocht met gestructureerde interviews. Uit beschrijvende statistiek bleek dat hogere stadia samengingen met het vaker hebben van angststoornissen twee jaar, vier jaar of zes jaar later. Het percentage aanwezigheid van angststoornissen bij een follow-up na twee jaar was bijvoorbeeld 2,7% voor de proefpersonen die oorspronkelijk in stadium 0 werden ingedeeld terwijl dit percentage 68,0% betrof voor de proefpersonen die oorspronkelijk in stadium 4B zaten.

De construct validiteit werd beoordeeld door verschillende klinische kenmerken op de baselinemeting te vergelijken tussen de verschillende stadia. Hierbij was van belang dat deze kenmerken niet waren gebruikt bij het toekennen van proefpersonen tot de verschillende stadia. De hypothese was dat hogere klinische stadia geassocieerd zijn met minder gunstige klinische parameters op baseline en dat deze associaties een lineaire trend zouden volgen langs elk van de stadia. Dit was inderdaad het geval: patiënten in hogere stadia hadden meer aanwezigheid van traumatische jeugdervaringen, een lagere leeftijd bij het optreden van de angststoornis, meer actuele psychiatrische behandelingen, een hogere ernst van de angstklachten, meer sociale en agorafobische vermijding en hogere mate van piekeren. Bovendien hadden patiënten in hogere stadia meer depressieve symptomen en een hogere mate van beperkingen. Niet-parametrische tests voor lineaire trends waren significant voor alle vooraf gedefinieerde validatoren na Bonferroni-correctie. Dit impliceert dat hogere klinische stadia geassocieerd waren met minder gunstige klinische kenmerken bij aanvang, wat wijst op een adequate constructvaliditeit.

Om de voorspellende validiteit van het model te evalueren, werden verschillende analyses uitgevoerd. Ten eerste werd het klinische stadiëringsmodel gerelateerd aan de aanwezigheid van DSM-IV-diagnoses na verloop van tijd. Er werden twee sets van *generalized estimating equations* (GEE) uitgevoerd om voor elk stadium te berekenen wat de kans was om ofwel een angststoornis of enigerlei psychiatrische stoornis te hebben op elke van de drie opeenvolgende tijdstippen (2 jaar, 4 jaar

en 6 jaar follow-up). De hypothese was dat hogere klinische stadia een grotere kans hadden op het hebben van angststoornis diagnoses en overige psychiatrische diagnoses op elke follow-up meting. Het bleek dat dit inderdaad het geval was: *odds-ratios* (OR) voor het hebben van een angststoornis bij de follow-up meting na 6 jaar was 11,8 (95% BI 8,39-16,6) voor stadium 4B afgezet tegen stadia 0-1B. De OR voor het hebben van enigerlei psychiatrische stoornis na 6 jaar follow-up was 10,7 (95% BI 7,70-15,0) voor stadium 4B afgezet tegen stadia 0-1B. Voor alle stadia waren de 6-jaars proporties van aanwezigheid van psychiatrische diagnosen lager in vergelijking tot de 2-jaars proporties. Voor de primaire uitkomstmaat, de aanwezigheid van angststoornissen, werd een lineaire trend over de baseline-stadia gevonden. De B-stadia bleken de grootste kans te hebben voor de aanwezigheid van enigerlei psychiatrische stoornis bij follow-up. Over het algemeen vertoonde ook dit model een significante lineaire trend over alle stadia. Deze bevindingen impliceren dat het risico op het hebben van een angststoornis bij de follow-up hoger is voor elke volgende fase, maar risico's op het hebben van enigerlei psychiatrische stoornis zijn het grootst bij B-stadia waarbij al psychiatrische comorbiditeit aanwezig was bij de baselinemeting. Als aanvulling op deze dichotome benadering werd de voorspellende validiteit ook beoordeeld met behulp van een dimensionele benadering. De klinische stadia bij de eerste meting werden gerelateerd aan follow-up metingen van de ernst van de angstklachten, de ernst van de depressieve klachten en de ernst van eventuele beperkingen door gebruik te maken van *linear mixed models* (LMM), waarmee ontbrekende gegevens nauwkeurig konden worden geschat. Dit leverde geschatte gemiddelden op voor elk stadium op alle follow-up metingen. Uit LMM-analyses bleek dat de ernst van de angst in de loop van de tijd geleidelijk afnam voor de meeste stadia, terwijl de rangorde die op de baseline meting aanwezig was tussen de verschillende stadia grotendeels behouden bleef. Stadia 2B, 3A en 4A vertoonden vergelijkbare maten van ernst van angstklachten over de tijd, terwijl stadia 3B en 4B de hoogste angstniveaus, de hoogste ernst van depressieve klachten en de hoogste mate van beperkingen lieten zien.

Samenvattend toonde dit hoofdstuk de eerste succesvolle poging om een stadiëringsmodel empirisch toe te passen op een cohort van angststoornispatiënten en controle personen. De resultaten laten zien dat het bestudeerde model voldoende construct- en voorspellende validiteit heeft, en daarmee een *evidence-based* stadiëringsmodel biedt voor gebruik in de klinische zorg bij patiënten met angststoornissen.

Het doel van **hoofdstuk 6** was om de bevindingen uit hoofdstuk 4 empirisch te toetsen. Hoofdstuk 4 leverde een aantal criteria op die gezamenlijk de definitie voor therapieresistentie bij angststoornissen (TR-AD) vormen. Het doel in dit hoofdstuk was om te toetsen of de aanwezigheid van deze criteria ook daadwerkelijk samengaat met slechte resultaten na behandeling. De hypothese hierbij was dat de aanwezigheid van de verschillende aspecten voor TR-AD inderdaad samenhangt met minder gunstige latere therapie-effecten. Een tweede doel van deze studie was om een dimensioneel meetinstrument te ontwikkelen door de verschillende TR-AD-criteria te integreren tot één maat. Dit dimensionele instrument zou dan de *mate* van TR-AD weergeven, wat een relevante aanvulling zou kunnen zijn op de dichotome benadering van TR-AD die in hoofdstuk 4 werd toegepast.

Voor deze studie werd uit 1305 NESDA-patiënten met angststoornissen op de baseline meting een steekproef van 679 patiënten genomen die aangaven psychiatrische behandelingen te hebben gehad tussen de baselinemeting en de meting tijdens de follow-up na 2 jaar. Alle criteria voor TR-AD die werden afgeleid uit het systematische review (hoofdstuk 4) werden op de baselinemeting beoordeeld. Deze TR-AD-criteria betroffen het aantal eerstelijns farmacologische behandelingen, aantal tweedelijns farmacologische behandelingen, aantal adequate psychotherapeutische behandelingen, ernst van angstklachten, aanwezigheid van functionele beperkingen, aanwezigheid van psychiatrische comorbiditeit en eerdere duur van angstklachten. Op basis van deze criteria werd een meetinstrument ontwikkeld. Om dit te doen, werd elk van de TR-AD-criteria gescoord in overeenstemming met de *Dutch Measure for quantification of Treatment Resistant Depression* (DM-TRD), wat een dimensioneel meetinstrument voor de mate van therapieresistentie bij depressies is. Op deze manier werd een meetinstrument gemaakt met een theoretische range van 2-23.

Op baseline was de gemiddelde score op het meetinstrument 10,8 ± 2,3. Bivariate logistische regressieanalyses werden gebruikt om individuele baseline TR-AD-criteria te relateren aan het behandelresultaat na twee jaar door te beoordelen in welke mate na twee jaar nog angststoornisdiagnoses aanwezig waren. Een hogere ernst van angstklachten op baseline was duidelijk geassocieerd met slechtere resultaten na twee jaar: de OR voor het hebben van een angststoornis na twee jaar was 6,48 (95% BI 3,29-12,8) in vergelijking met proefpersonen met lage ernst van angstklachten bij baseline. Hoge niveaus van functionele beperkingen waren ook geassocieerd met slechte resultaten na twee jaar follow-up: OR=2,90 (95%

BI 1,51-5,63). Dezelfde positieve associaties werden gevonden voor de aanwezigheid van psychiatrische comorbiditeit (OR=1,73, 95% CI 1,25-2,39) en lange duur van angstsymptomen (OR=2,79, 95% CI 1,94-4,03 versus korte duur). Het eerdere aantal eerstelijns farmacotherapeutische behandelingen was niet geassocieerd met uitkomsten na twee jaar (OR=1,10, 95% BI 0,87-1,39). Evenmin was het aantal eerdere tweedelijns farmacotherapeutische behandelingen (OR=1,39, 95% BI 0,91-2,12), noch het aantal eerdere psychotherapeutische behandelingen (OR=1,11, 95% BI 0,73-1,69) geassocieerd met slechtere uitkomsten na 2 jaar. Hogere scores op het dimensionele meetinstrument in zijn geheel waren duidelijk geassocieerd met slechtere behandelresultaten na twee jaar follow-up: OR=1,29, 95% BI 1,20-1,39, wat aangeeft dat de kans op het aanwezig blijven van angststoornissen na twee jaar met 1,29 hoger uitvalt voor elk punt verschil op het meetinstrument.

Tot slot werden de klinimetrische eigenschappen van dit meetinstrument beoordeeld. Met behulp van de Youden-index werd de meest efficiënte afkapwaarde voor het meetinstrument vastgesteld op 11 punten of hoger. Met deze afkapwaarde was de sensitiviteit voor het voorspellen van aanwezig blijven van angststoornissen na een periode van 2 jaar waarin patiënten behandeld werden 0,70, terwijl de specificiteit 0,57 was. Hieruit werd de positief voorspellende waarde berekend op 0,68 en de negatief voorspellende waarde op 0,60. De oppervlakte onder de curve (AUC) was 0,66. De kracht van deze voorspellingen kan als matig worden beschouwd. Echter, vanwege het gebrek aan een gouden standaard waarmee de behandelresultaten bij angststoornissen nauwkeurig kunnen worden voorspeld geeft het hier gepresenteerde meetinstrument momenteel de beste nauwkeurigheid. Het gebruik van dit meetinstrument kan bijdragen aan het effectief selecteren van behandelstrategieën bij individuele patiënten. Voordat dit in de praktijk kan worden toegepast zou deze studie echter eerst moeten worden gerepliceerd in een andere steekproef om de generaliseerbaarheid van deze bevindingen te beoordelen.

Het doel van **hoofdstuk 7** was om met behulp van *machine learning* methodes op basis van een grote verscheidenheid aan data een voorspellingsmodel te ontwikkelen voor het longitudinale beloop bij patiënten met angststoornissen. Van veel risicofactoren is eerder al aangetoond dat ze samenhangen met het optreden van angststoornissen of met het longitudinale beloop van angststoornissen. Geen enkele risicofactor heeft echter zodanige adequate voorspellende eigenschappen

dat ze in de klinische praktijk gebruikt kunnen worden voor het geven van een betrouwbare prognose. Voorheen werden deze voorspellers veelal afzonderlijk onderzocht, in plaats van als onderdeel van één geïntegreerd model. Mogelijk kan het combineren van een veelheid van deze voorspellers in één model resulteren in betere voorspellende waarde. Machine learning methodes zijn bij uitstek geschikt voor het maken van voorspellingen op basis van een grote verscheidenheid aan factoren, omdat ze een data-gestuurde aanpak kunnen gebruiken om de meest relevante factoren te identificeren en te selecteren. De hypothese was dat het combineren van veel voorspellers in één grootschalig model zou resulteren in een model met goede voorspellende waarde wat betreft het longitudinale beloop bij angststoornissen.

De huidige studie werd uitgevoerd bij 887 patiënten met een angststoornis uit de NESDA. Voor dit onderzoek zijn verschillende potentiële voorspellers uit vijf domeinen gebruikt. De vijf domeinen zijn klinische variabelen, psychologische variabelen, biologische variabelen, sociodemografische variabelen en leefstijlvariabelen. De onderzochte classificaties waren tweeledig: ten eerste herstel van angststoornissen na twee jaar follow-up. Ten tweede, herstel van alle veelvoorkomende psychiatrische stoornissen (*common mental disorders,* CMD's) na een follow-up van twee jaar. De angststoornissen die onderzocht werden waren de gegeneraliseerde angststoornis, de paniekstoornis, agorafobie en de sociale angststoornis. CMD's werden gedefinieerd als ofwel een angststoornis, ofwel een depressieve stoornis, ofwel dysthymie, ofwel alcoholafhankelijkheid. Voor de tweede classificatie is er dus sprake van herstel wanneer bij follow-up geen van deze stoornissen kon worden gediagnosticeerd. Om de modellen te bouwen, werden *Random Forest Classifiers* (RFC's) gebruikt. Een RFC is opgebouwd als een combinatie van meerdere beslisbomen, die worden getraind door willekeurige subsets van variabelen en patiënten te gebruiken voor elke beslisboom. In dit onderzoek bestond elke RFC uit 1.000 beslisbomen. Op baseline werden 651 individuele voorspellers op itemniveau geselecteerd uit de vijf domeinen (klinisch, psychologisch, biologisch, sociodemografisch, leefstijl). Er bleven 569 items over nadat de items met te veel missing data waren verwijderd. De RFC's werden gebouwd met behulp van deze 569 items en met behulp van een 10x10 cross-validatie op trainingssets (90% van de proefpersonen) en testsets (de resterende 10% van de proefpersonen). Het algoritme werd

getraind op de combinatie van alle predictordomeinen, maar ook op elk verschillend domein afzonderlijk, om te zien in welke mate de afzonderlijke domeinen bijdragend zijn aan de betrouwbaarheid van de voorspellingen.

Na twee jaar follow-up herstelden 484 patiënten (54,6%) van hun angststoornis en hadden 362 patiënten (40,8%) geen CMD. De prestatie van de RFC's bij het voorspellen van herstel van angststoornissen na twee jaar follow-up was matig: de *area under the receiver-operator curve* (AUC) voor het volledige model bedroeg 0,67. Wanneer we de individuele domeinen onderzochten varieerde de AUC van 0,49 (leefstijl domein) tot 0,67 (klinische domein). De enige domeinen met statistisch significante voorspellingen over twee jaar waren het klinische domein (AUC=0.67) en het psychologische domein (AUC=0.65). Met betrekking tot de tweede uitkomstmaat (herstel van CMD's) leken de voorspellingen iets nauwkeuriger te zijn. De combinatie van alle domeinen leverde een AUC van 0,70 op. Bij de individuele domeinen varieerde de AUC voor herstel van CMD's van 0,53 (leefstijl domein) tot 0,70 (klinische domein). In deze analyse leverde het klinische domein (AUC=0,70), het psychologische domein (AUC=0,67) en het sociodemografische domein (AUC=0,65) significante voorspellingen op.

Post-hoc analyses werden uitgevoerd om de individuele items te identificeren die in hoge mate hebben bijgedragen aan de voorspellingen. Dit werd gedaan om meer inzicht te krijgen in het relatieve belang van elk van de voorspellers voor het longitudinaal beloop bij angststoornissen, in samenhang met de andere variabelen. Bij het voorspellen van herstel van angststoornissen werden 17 items geselecteerd in meer dan 50% van de RFC's, waarmee werd voldaan aan het criterium voor consistente selectie. Deze voorspellers waren afkomstig uit het klinische en psychologische domein en waren gerelateerd aan angstreacties, meestal aan angstige lichamelijke reacties. Bij het voorspellen van herstel van CMD's werden 48 variabelen consistent (>50% van de RFCs) geselecteerd. Eén geselecteerde voorspeller was afkomstig uit het sociodemografische domein, alle andere uit de klinische en psychologische domeinen. Behalve items die betrekking hadden op uitingen van angst werden bij deze classificatie consequent veel stemmingsgerelateerde items geselecteerd. In vergelijking met de variabelen die consequent werden geselecteerd bij het voorspellen van herstel van angststoornissen, leek het herstel van CMD's meer afhankelijk te zijn van variabelen die gerelateerd zijn aan uitingen van depressie. De prestaties van de voorspellingsmodellen in dit hoofdstuk zijn nog onvoldoende om deze routinematig in de klinische praktijk te gebruiken.

Bovendien zou implementatie hiervan in de klinische praktijk door het grote aantal metingen dat nodig is om dit model toe te passen onhaalbaar zijn. Dat de huidige redelijke voorspellende waardes gehaald werden ondanks de lange termijn van follow-up van twee jaar is echter veelbelovend. Dit onderzoek toont duidelijk het potentieel van machine learning methodes voor het verbeteren van voorspellingsmodellen in een complex vakgebied als de psychiatrie.

Dankwoord

Na vele analyses, woorden, zinnen, paragrafen en overpeinzingen wil ik u welkom heten bij mijn laatste woorden, zinnen en paragrafen in dit dankwoord. Het voelt gek om dit te schrijven en me te realiseren dat dit echt de allerlaatste woorden van dit proefschrift zijn. En feitelijk ligt de druk er flink op aangezien het dankwoord toch meestal het meest-gelezen stuk van een thesis blijkt te zijn. Dus daar gaan we. Dit lange en bij vlagen moeilijke traject had ik niet kunnen afronden zonder veel hulp. Bedankt!

Allereerst Ton. Naast mijn opleider was je ook mijn promotor en vanuit beide rollen heb je me altijd alle mogelijkheden gegund om me te ontwikkelen op een manier die bij mij paste. Je maakte deze promotie mogelijk door de inmiddels eigenlijk al doodverklaarde AGIKO-constructie met kunst- en vliegwerk levend te houden. Je bezit een Obama-achtige positivisme met een onweerstaanbare '*yes we can*' mentaliteit. Bij meerdere van mijn papers ontstonden er momenten waarop het hele proces wat dreigde vast te lopen en dan was één afspraak bij jou telkens voldoende om weer met een hernieuwd optimisme (en een hernieuwde to do lijst) terug te keren. Je didactische kwaliteiten staan buiten kijf en ik ben blij dat ik mijn voordrachten bij je kon oefenen en dat ik je feedback kon gebruiken. Regelmatig adviseerde je mij om mijn publiek niet te overschatten, wat bij mij een soort onbestemd maar zelfverzekerd gevoel opriep waarmee mijn voordrachten steeds leuker werden om te geven. Als laatste ben je een gangmaker met je humor en aanstekelijke lach en heb je menig lunch op de Oldenaller en menig opleidingsuitje op een manier positief gekleurd zoals alleen jij dat kan. Bedankt dat je met verve meerdere petten droeg!

Vervolgens Brenda. Ik weet nog goed dat ik aan het begin van mijn tijd als onderzoeker eens googlede naar jouw publicaties en dat ik steil achterover sloeg van wat ik tegenkwam. Het werd me direct duidelijk dat je één van de belangrijkste onderzoekers binnen de psychiatrie wereldwijd bent en het maakte me trots dat ik je mijn promotor mag noemen. Het was soms hard werken om het tempo van jouw feedback en het tempo van je argumentatie te volgen en te integreren. Ik leerde van je om niet voor de makkelijke weg te kiezen maar om telkens ambitieus en kritisch te blijven ten aanzien van mijn eigen werk. Je tilde mijn papers daarmee keer op keer naar een hoger niveau. Het was erg prettig met je samenwerken, bedankt!

Dan Neeltje. Als copromotor stond je het meest dicht bij mij en mijn projecten. Ik vond jou het perfecte verbindingsstuk tussen Ton, Brenda en mijzelf. Als voormalig protégé van hen beide is het niet vreemd dat ik in jou kwaliteiten van zowel Ton als Brenda herken. Je was de schakel die jullie tot een erg prettig en efficiënt team maakte. Je was altijd beschikbaar voor overleg over grote en kleine onderwerpen. Daarnaast heb je onmiskenbaar je eigen stijl: positief, kritisch, benaderbaar en respectvol. Je bent een held in het strategisch formuleren van tegenwerpingen richting onredelijke reviewers die je ontdeed van de Hollandse directheid waar ik me in het begin nog regelmatig schuldig aan maakte. Jouw mailtjes naar mij begonnen steevast met zinnen als "Heel goed! Het paper begint al richting de laatste fase te gaan", waarvan ik na verloop van tijd steeds beter begon te begrijpen dat je eigenlijk bedoelde "heel leuk begin, maar er moet nog wel een hele boel gebeuren". Ik denk dat je de allerlaatste persoon op aarde was om de overstap naar digitaal feedback geven te maken. Potloodfabrikanten over de hele wereld hebben met jou een gouden tijd achter de rug want je feedback kwam steevast in redelijk leesbare krabbels in de kantlijn van mijn printjes te staan. Bedankt voor al je krabbels en wijze raad!

Verder wil ik de overige medeauteurs bedanken voor hun bijdrages aan onze gezamenlijk papers. Aernout Beek, Annette Boenink en Jan Smit bedankt voor jullie bijdrage aan het eerste harthulp paper. Adriaan Hoogendoorn, bedankt voor je hulp en creativiteit bij de lastige statistiek van het stadiërings paper en je auteurloze hulp bij meerdere andere papers. Guido Wetzer en Jurriaan Gehrels, bedankt voor jullie inzet bij het doorploegen van de immense berg papers bij het systematische review. Erik Giltay, Robert Schoevers en Guido van Wingen, bedankt voor jullie bijdrage aan het machine learning paper. And thanks Paul Zhutovsky for a very pleasant cooperation as shared first authors on the machine learning paper. We were a pretty effective team if you ask me.

Vervolgens wil ik graag de vele respondenten die hebben deelgenomen aan de NESDA studie bedanken. Ik realiseer me dat het deelnemen aan een studie zoals NESDA veel vergt van respondenten, zeker wanneer er sprake is van een actuele depressie of angststoornis. En dan ook nog eens een studie als NESDA met zijn zeer uitgebreide dataverzameling. Bedankt voor jullie inzet! Ik ben er van overtuigd dat jullie deelname voor veel andere huidige en toekomstige patiënten belangrijke winst oplevert.

Dan kan het niet anders dan dat kamer D2.05 als volgende aan de beurt is. Ilja, Ruth, Sjors en Lucas, in wisselende samenstellingen vormden wij over de jaren gezamenlijk een gerenommeerde GGZ/politieke denktank. Later ook nog aangevuld met AMC spion Angela. Eerst in de penthouse onder de naam D2.05, later in een meer bescheiden en donker hoekje van het gebouw. Ilja, ontzettend bedankt voor al jouw hulp bij mijn project en de projecten van iedereen, en dat je dan zelf ondertussen ook nog gepromoveerd bent, onvoorstelbaar. Sjors bedankt dat ik mocht meedelen in al je crackers, pindakaas, gemengde noten en je verhalen. Je eigen uitspraak "neem ruim" typeert je perfect. Ruth, bedankt voor je inspirerende voorbeeld in het bijhouden van email, ongeëvenaard. Angela, deed jij eigenlijk wel aan onderzoek, of zat je alleen maar aan zee? Lucas, bedankt voor de inspirerende gesprekken en gezellig dat je altijd even mee lunchte.

Ook de andere junior onderzoekers op de Oldenaller hebben de onderzoeksjaren een heel stuk gezelliger gemaakt met veel lunches, meetings, borrels en een feestje hier en daar. Als groentje met Dóra, Josine en de rest pull-ups doen in het park. Later onder andere met Trees, Esther, Laura, Richard et al. Bedankt allemaal! Hopelijk komen we elkaar nog tegen op toekomstige werkvloeren of dansvloeren.

Verder waren de opleidingsjaren voor mij een welkome afwisseling van de onderzoeksjaren, en andersom. Doordat het combineren van de opleiding met het onderzoek voor een lekker lang traject heeft gezorgd heb ik extra veel mede aiossen mogen ontmoeten. Bedankt voor vele vrijdagmiddag borrels, prachtige weekenden weg, leuke voorjaarscongressen en gewoon een gouden tijd. Bedankt Noach, Linda, Saskia, Afra, Ruud, Hans, Ellemijn, Esther, Peter, Marc en Frits en nog vele anderen, en bedankt AGIKO-voorgangers en voorbeelden Sonja, Hans, Wouter en Flora.

Ook bedankt aan onze intervisiegroep "Duur & Delicious" voor alweer acht jaar levensintervisie, Thais eten, sterke verhalen, vermijdende input en confronterende adviezen! Stella, Laura, Kathelijne, Lidwien, Larissa, Doeke en Martijn; dat er nog maar vele mooie avonden mogen volgen.

Omdat de boog niet altijd gespannen kan zijn ook een shout-out naar alle belangrijke vrienden. Ruud, Erik, Erik, Siert, Jorrit, SyMa, Johan, Con, Kris-Jan, Sander, Steven, Casper en Ruben plus alle respectievelijke aanhang: bedankt!

Bedankt Martin van den Berg voor de illustraties in dit proefschrift. Mijn eerste en jouw laatste jaar als psychiater vielen samen in Sneek. Bedankt voor je raad en relativeringsvermogen!

Een extra bedankje voor Peter en Erik, bedankt dat jullie me als mijn paranimfen willen vergezellen naar de grote, intimiderende aula. Peblo, we begonnen samen in de Valeriuskliniek, jij als semi-arts, ik als aios. Later troffen we elkaar telkens weer, zelfs op het voetbalveld en in de derde helft. De vriendschap groeide als vanzelfsprekend. Bedankt voor je openheid, droge humor, warmte en bescheidenheid. En mocht ik omvallen in de aula verwacht ik veel van je als plaatsvervanger. Erik, bedankt dat jij de Amsterdamse groep wil vertegenwoordigen! We gaan inmiddels zo'n achttien jaar terug. In de beginjaren waren we een roemrucht DJ-duo. We delen een prachtige kerstdiner traditie met de rest van de eerder genoemde mannen. Het is mooi om te zien hoe je in die jaren altijd precies jezelf bent gebleven ondanks heel uiteenlopende paden die je insloeg. Je maakte uitstapjes naar andere werelddelen en andere carrièrepaden en je kwam telkens terug. Een vriendschap voor het leven.

Aukje, Liesbeth, Lydia. Het is toch wat, drie schoonmoeders… gelukkig is het in het echt lang niet zo erg als op papier. Fijn dat TWARLL bestaat! Dat is niet vanzelfsprekend maar voelt wel zo. Bedankt dat ik bij jullie kan lachen en dat we veel kunnen delen. En Liesbeth, je zal wel trots zijn op je schoonzoon psycholoog, of was het nou psychiater? Rosan, bedankt voor je no nonsense houding en je eerlijkheid! Leuk om je te mogen zien groeien de laatste jaren.

Bedankt JAF, jullie zijn al mijn hele leven een stabiele twee-eenheid waardoor wij kinders alle ruimte hebben voor onze ontwikkeling. Allebaas, ondanks dat je je werk niet mee naar huis nam was je altijd een grote inspiratie voor me om ook het doktersvak te kiezen. Je was je tijd ver vooruit als mannelijke arts met een part-time dienstverband en een privéleven. Je was vast de enige medisch specialist die in die jaren nachtbraakte met Age of Empires. "Haa-yee-yoo-haa-hoo, homme-nahhh!" Je bent van alle markten thuis en ik kan je dus nu ook bedanken voor je steun op vele fronten. Dr. Moekes, bedankt voor de vele kleine zetjes de goede richting op die ik zelf pas achteraf opmerkte. Je bent een echte problemenoplosser. En soms zelfs een problemenvoorkómer. Je hebt bovendien toch een beetje de weg geplaveid voor Jouke en mij in de wetenschappelijke wereld. Het waren een paar spannende jaren qua gezondheid, gelukkig kunnen we die nu achter ons laten en heel knap hoe je die jaren optimistisch hebt volgebracht.

Bedankt voor al je zichtbare en onzichtbare acties! Jouke, mooi om onze geneeskunde studententijd te kunnen delen en tof dat we altijd welkom zijn bij elkaar en elkaars vrienden. Jammer dat mijn knieën niet meer in staat zijn voor één-tweetjes op het voetbalveld maar Pok en Tok weten elkaar daarbuiten gelukkig ook uitstekend te vinden. Mijke, Kippie, bedankt voor je scherpzinnigheid en je loyaliteit! Je hebt een heerlijk gezin waar ik me erg mee verbonden voel. Bij jullie kleintjes is de stress van werk opeens heel ver weg. Ik voel me gezegend met jullie allemaal als gezin. Daarbij mogen natuurlijk ook Magic Wierc en die Karo niet ontbreken. Wierc, bedankt voor je rust en humor. Je bent een stille kracht op vele manieren.

Als laatste, bedankt liefste Tamar, TEA. Je bent altijd geduldig, begripvol en geïnteresseerd wanneer er werk gerelateerde obstakels waren. Bedankt voor je creativiteit en je vermogen om altijd een oplossing te vinden waarmee de gezamenlijkheid en gelijkwaardigheid voorop staat. Je maakte met je aanwezigheid van de weekendjes dienst in Sneek een feestje. Het samen aan onze papers werken in het 'prachtige' Amicitia hotel voelde als een enorme life-hack. Je liet me mijn presentaties bij je oefenen en was mijn grootste fan en gaf ook op subtiele wijze kritische feedback. Bedankt voor het herontdekken van je lievelingslied Ademnood en van het woord 'fijn'. We kunnen elkaar in meerdere versnellingen goed vinden en ik kijk er naar uit om nog vele jaren samen onze levens te kunnen delen.

Curriculum vitae

Wicher Alle Bokma was born on December 31st 1983 in Amersfoort, the Netherlands. At the age of five he moved to Heemstede where he later attended high school at College Hageveld, from which he graduated in 2002. From 2002 to 2005 he studied psychology at the Vrije Universiteit, from which he graduated with a Bachelor's degree. From 2005 to 2012 he studied Medicine at the same university, during which he did scientific research into the mental health of young adult postoperative transgender subjects under supervision of child and adolescent psychiatrist dr. A.L.C. de Vries at Amsterdam UMC, locatie VUmc. In the final year of his master's degree he did an internship ('semi-arts') at the psychiatric department of the Spaarne Gasthuis (formerly Kennemer Gasthuis) in which he developed his interest in general hospital psychiatry. After graduating as a MD he continued working at the Spaarne Gasthuis as an general hospital psychiatry ANIOS and subsequently got accepted for a psychiatry residency at GGZ inGeest with prof. dr. Anton van Balkom. After a couple of months of traveling across southeast Asia and awaiting the start of his psychiatry residency he started as a junior scientist at GGZ inGeest in June 2013. The foundation for the start of a PhD-trajectory in combination with the residency programme which started in October 2013 was laid in these couple of months during which he started working on the CED project that was included as the second chapter in this thesis. From October 2014 onwards, he continued combining clinical work and scientific work under the supervision of Prof. dr. Ton van Balkom, Prof. dr. Brenda Penninx and dr. Neeltje Batelaan, which resulted in this thesis. During these years, Wicher finished a masters degree to become an epidemiologist in 2018. During his residency he treated patients at the medium/high care inpatient facility (formerly Valeriuskliniek, now De Nieuwe Valerius) under supervision of Nikander Ruhl, a high/intensive care inpatient facility (supervised by Annet Ferwerda), the anxiety- and obsessive compulsive disorders outpatient facility (supervised by Prof. dr. Gerben Meynen) and the FACT outpatient facility located at the Hilligaertstraat (supervised by Erika Verkruissen). Further training in general hospital psychiatry was done at Amsterdam UMC (locatie VUmc) under supervision of Klaas-Jan Nauta and Hansje Heller. Wicher chose a period of somatic training with the internal medicine department of Amsterdam UMC (locatie VUmc) in which he focussed on geriatric internal medicine and dementias under supervision of Gooke Lagaay. Finally, he worked as a last year psychiatry resident at the crisis psychiatric department of Arkin (spoedeisende psychiatrie) under supervision of dr. Julia Meijer.

After finishing the residency programme in April 2020 he started working as a general hospital psychiatrist at the Antonius Hospital in Sneek. He currently lives in Amsterdam with his girlfriend Tamar.

Publication list

First author publications (published)

Bokma, W.A., Batelaan, N.M., Beek, A.M., Boenink, A.D., Smit, J.H., van Balkom, A.L.J.M. (2015). Feasibility and outcome of the implementation of a screening program for panic disorder in noncardiac chest pain patients in cardiac emergency department routine care. *General Hospital Psychiatry. 37: 485–487.*

Bokma, W.A., Batelaan, N.M., van Balkom, A.L.J.M., Penninx, B.W.J.H. (2017). Impact of Anxiety and/or Depressive Disorders and Chronic Somatic Diseases on disability and work impairment. *Journal of Psychosomatic Research. 94: 1–16.*

Bokma, W.A., Batelaan, N.M., Hoogendoorn, A.W., Penninx, B.W.J.H., van Balkom, A.L.J.M. (2020). A clinical staging approach to improving diagnostics in anxiety disorders: Is it the way to go? *Australian and New Zealand Journal of Psychiatry. 54: 173–184.*

Bokma, W.A., Batelaan, N.M., Penninx, B.W.J.H., van Balkom, A.L.J.M. (2021). Evaluating a dimensional approach to treatment resistance in anxiety disorders: A two-year follow-up study, *Journal of Affective Disorders Reports. 4: 00139.*

Bokma, W.A., Wetzer, G.A.A.M., Gehrels, J.B., Penninx, B.W.J.H., Batelaan, N.M., van Balkom, A.L.J.M. (2019). Aligning the many definitions of treatment resistance in Anxiety Disorders: a systematic review., *Depression and Anxiety. 36: 801–812.*

Bokma, W.A.*, Zhutovsky, P.*, Giltay, E.J., Schoevers, R.A., Penninx, B.W.J.H., van Balkom, A.L.J.M., Batelaan, N.M.*, van Wingen, G.A.* (2020). Predicting the naturalistic course in anxiety disorders using clinical and biological markers: a machine learning approach. *Psychological Medicine. 1–11. * both authors contributed equally.*

Co-author publications (in preparation)

Seldenrijk, A., **Bokma, W.A.**, Jacobs, G., van Balkom, A.J.L.M., Ruhé, H.G.*, Batelaan, N.M.* (in preparation). The effects of tryptophan depletion on mood: an updated systematic review with meta-analysis and dose-response investigation.

*both authors contributed equally

Dissertation series

Department of Psychiatry, Amsterdam University Medical Centers

N.M. (Neeltje) Batelaan (2010). Panic and Public Health: Diagnosis, Prognosis and Consequences. Vrije Universiteit Amsterdam. ISBN: 978-90-8659-411-5.

G.E. (Gideon) Anholt (2010). Obsessive-Compulsive Disorder: Spectrum Theory and Issues in Measurement. Vrije Universiteit Amsterdam.

N. (Nicole) Vogelzangs (2010). Depression & Metabolic Syndrome. Vrije Universiteit Amsterdam. ISBN: 978-90-8659-447-4.

C.M.M. (Carmilla) Licht (2010). Autonomic Nervous System Functioning in Major Depression and Anxiety Disorders. Vrije Universiteit Amsterdam. ISBN: 978-90-8659-487-0.

S.A. (Sophie) Vreeburg (2010). Hypothalamic-Pituitary-Adrenal Axis Activity in Depressive and Anxiety Disorders. Vrije Universiteit Amsterdam. ISBN: 978-90-8659-491-7.

S.N.T.M. (Sigfried) Schouws (2011). Cognitive Impairment in Older Persons with Bipolar Disorder. Vrije Universiteit Amsterdam. ISBN: 978-90-9025-904-8.

P.L. (Peter) Remijnse (2011). Cognitive Flexibility in Obsessive-Compulsive Disorder and Major Depression – Functional Neuroimaging Studies on Reversal Learning and Task Switching. Vrije Universiteit Amsterdam. ISBN: 978-90-6464-449-8.

S.P. (Saskia) Wolfensberger (2011). Functional, Structural, and Molecular Imaging of the Risk for Anxiety and Depression. Vrije Universiteit Amsterdam. ISBN: 978-90-8659-536-5.

J.E. (Jenneke) Wiersma (2011). Psychological Characteristics and Treatment of Chronic Depression. Vrije Universiteit Amsterdam. ISBN: 978-94-9121-150-8.

P.D. (Paul David) Meesters (2011). Schizophrenia in Later Life. Studies on Prevalence, Phenomenology and Care Needs (SOUL Study). Vrije Universiteit Amsterdam. ISBN: 978-90-8659-563-1.

R. (Ritsaert) Lieverse (2011). Chronobiopsychosocial Perspectives of Old Age Major Depression: a Randomized Placebo Controlled Trial with Bright Light. Vrije Universiteit Amsterdam. ISBN: 978-90-8570-858-2.

A. (Adrie) Seldenrijk (2011). Depression, Anxiety and Subclinical Cardiovascular Disease. Vrije Universiteit Amsterdam. ISBN: 978-94-6191-052-3.

Y. (Yuri) Milaneschi (2012). Biological Aspects of Late-life Depression. Vrije Universiteit Amsterdam. ISBN: 978-90-8659-608-9.

L. (Lynn) Boschloo (2012). The Co-occurrence of Depression and Anxiety with Alcohol Use Disorders. Vrije Universiteit Amsterdam. ISBN: 978-94-6191-327-2.

D. (Didi) Rhebergen (2012). Insight into the heterogeneity of depressive disorders. Vrije Universiteit Amsterdam. ISBN: 978-94-6191-387-6.

T.M. (Michiel) van den Boogaard (2012). The Negotiated Approach in the Treatment of Depressive Disorders: the impact on patient-treatment compatibility and outcome. Vrije Universiteit Amsterdam. ISBN: 978-90-8891-495-9.

M. (Marjon) Nadort (2012). The implementation of outpatient schema therapy for borderline personality disorder in regular mental healthcare. Vrije Universiteit Amsterdam. ISBN: 978-94-6191-463-7.

U. (Ursula) Klumpers (2013). Neuroreceptor imaging of mood disorder related systems. Vrije Universiteit Amsterdam. ISBN: 978-94-6191-575-7.

E. (Ethy) Dorrepaal (2013). Before and beyond. Stabilizing Group treatment for Complex posttraumatic stress disorder related to child abuse based on psycho-education and cognitive behavioral therapy. Vrije Universiteit Amsterdam. ISBN: 978-94-6191-601-3.

K. (Kathleen) Thomaes (2013). Child abuse and recovery. Brain structure and function in child abuse related complex posttraumatic stress disorder and effects of treatment. Vrije Universiteit Amsterdam. ISBN: 978-94-6191-600-6.

K.M.L. (Klaas) Huijbregts (2013). Effectiveness and cost-effectiveness of the implementation of a collaborative care model for depressive patients in primary care. Vrije Universiteit Amsterdam. ISBN: 978-90-9027404-1.

T.O. (Tessa) van den Beukel (2013). Ethnic differences in survival on dialysis in Europe. The role of demographic, clinical and psychosocial factors. Vrije Universiteit Amsterdam. ISBN: 978-94-6108410-1.

A. (Agnes) Schrier (2013). Depression and anxiety in migrants in the Netherlands. Population studies on diagnosis and risk factors. Vrije Universiteit Amsterdam. ISBN: 978-94-6191-719-5.

B. (Barbara) Stringer (2013). Collaborative Care for patients with severe personality disorders. Challenges for the nursing profession. Vrije Universiteit Amsterdam. ISBN: 978-94-6191-809-3.

C.M. (Caroline) Sonnenberg (2013). Late life depression: sex differences in clinical presentation and medication use. Vrije Universiteit Amsterdam. ISBN: 978-94-6191-866-6.

Z. (Zsuzsika) Sjoerds (2013). Alcohol dependence across the brain: from vulnerability to compulsive drinking. Vrije Universiteit Amsterdam. ISBN: 978-90-8891-695-3.

V.J.A. (Victor) Buwalda (2013). Routine Outcome Monitoring in Dutch Psychiatry: Measurement, Instruments, Implementation and Outcome. Vrije Universiteit Amsterdam. ISBN: 978-94-6191-905-2.

J.G. (Josine) van Mill (2013). Sleep, depression and anxiety: an epidemiological perspective. Vrije Universiteit Amsterdam. ISBN: 978-94-6108-525-2.

S. (Saskia) Woudstra (2013). Framing depression in a SN[a]Pshot: Imaging risk factors of MDD. Vrije Universiteit Amsterdam. ISBN: 978-90-8891-751-6.

N.C.M. (Nicole) Korten (2014). Stress, depression and cognition across the lifespan. Vrije Universiteit Amsterdam. ISBN: 978-94-6108-748-5.

M.K. (Maarten) van Dijk (2014). Applicability and effectiveness of the Dutch Multidisciplinary Guidelines for the treatment of Anxiety Disorders in everyday clinical practice. Vrije Universiteit Amsterdam. ISBN: 978-94-92096-00-5.

I.M.J. (Ilse) van Beljouw (2015). Need for Help for Depressive Symptoms from Older Persons Perspectives: The Implementation of an Outreaching Intervention Programme. Vrije Universiteit Amsterdam. ISBN: 978-94-6259-496-8.

A.M.J. (Annemarie) Braamse (2015). Psychological aspects of hematopoietic stem cell transplantation in patients with hematological malignancies. Vrije Universiteit Amsterdam. ISBN: 978-94-6259-594-1.

A. (Annelies) van Loon (2015). The role of ethnicity in access to care and treatment of outpatients with depression and/or anxiety disorders in specialised care in Amsterdam the Netherlands. Vrije Universiteit Amsterdam. ISBN: 978-94-90791-34-6.

C. (Chris) Vriend (2015). (Dis)inhibition: imaging neuropsychiatry in Parkinson's disease. Vrije Universiteit Amsterdam. ISBN: 978-94-6295-115-0.

A.M. (Andrea) Ruissen (2015). Patient competence in obsessive compulsive disorder. An empirical ethical study. Vrije Universiteit Amsterdam. ISBN: 978-90-6464-856-4.

H.M.M. (Henny) Sinnema (2015). Tailored interventions to implement guideline recommendations for patients with anxiety and depression in general practice. Vrije Universiteit Amsterdam. ISBN: 978-94-6169-653-3.

T.Y.G. (Nienke) van der Voort (2015). Collaborative Care for patients with bipolar disorder. Vrije Universiteit Amsterdam. ISBN: 978-94-6259-646-7.

W. (Wim) Houtjes (2015). Needs of elderly people with late-life depression; challenges for care improvement. Vrije Universiteit Amsterdam. ISBN: 978-94-6108-985-4.

M. (Marieke) Michielsen (2015). ADHD in older adults. Prevalence and psychosocial functioning. Vrije Universiteit Amsterdam. ISBN: 978-90-5383-132-8.

S.M. (Sanne) Hendriks (2016). Anxiety disorders. Symptom dimensions, course and disability. Vrije Universiteit Amsterdam. ISBN: 978-94-6259-963-5.

E.J. (Evert) Semeijn (2016). ADHD in older adults; diagnosis, physical health and mental functioning. Vrije Universiteit Amsterdam. ISBN: 978-94-6233-190-7.

N. (Noera) Kieviet (2016). Neonatal symptoms after exposure to antidepressants in utero. Vrije Universiteit Amsterdam. ISBN: 978-94-6169-794-3.

W.L. (Bert) Loosman (2016). Depressive and anxiety symptoms in Dutch chronic kidney disease patients. Vrije Universiteit Amsterdam. ISBN: 987-94-6169-793-6.

E. (Ellen) Generaal (2016). Chronic pain: the role of biological and psychosocial factors. Vrije Universiteit Amsterdam. ISBN: 978-94-028-0032-6.

D. (Dóra) Révész (2016). The interplay between biological stress and cellular aging: An epidemiological perspective. Vrije Universiteit Amsterdam. ISBN: 978-94-028-0109-5.

F.E. (Froukje) de Vries (2016). The obsessive-compulsive and tic-related brain. Vrije Universiteit Amsterdam. ISBN: 978-94-629-5481-6.

J.E. (Josine) Verhoeven (2016). Depression, anxiety and cellular aging: does feeling blue make you grey? Vrije Universiteit Amsterdam. ISBN: 978-94-028-0069-2.

A.M. (Marijke) van Haeften-van Dijk (2016). Social participation and quality of life in dementia: Implementation and effects of interventions using social participation as strategy to improve quality of life of people with dementia and their carers. Vrije Universiteit Amsterdam. ISBN: 978-94-6233-341-3.

P.M. (Pierre) Bet (2016). Pharmacoepidemiology of depression and anxiety. Vrije Universiteit Amsterdam. ISBN: 978-94-6299-388-4.

M.L. (Mardien) Oudega (2016). Late life depression, brain characteristics and response to ECT. Vrije Universiteit Amsterdam. ISBN: 978-94-6295-396-3.

H.A.D. (Henny) Visser (2016). Obsessive-Compulsive Disorder; Unresolved Issues, Poor Insight and Psychological Treatment. Vrije Universiteit Amsterdam. ISBN: 978-94-028-0259-7.

E.C. (Eva) Verbeek (2017). Fine mapping candidate genes for major depressive disorder: Connecting the dots. Vrije Universiteit Amsterdam. ISBN: 978-94-028-0439-3.

S. (Stella) de Wit (2017). In de loop: Neuroimaging Cognitive Control in Obsessive-Compulsive Disorder. Vrije Universiteit Amsterdam. ISBN: 978-90-5383-225 7.

W.J. (Wouter) Peyrot (2017). The complex link between genetic effects and environment in depression. Vrije Universiteit Amsterdam. ISBN: 978-94-6182-735-7.

R.E. (Rosa) Boeschoten (2017). Depression in Multiple Sclerosis: Prevalence Profile and Treatment. Vrije Universiteit Amsterdam. ISBN: 978-94-028-0474-4.

G.L.G. (Gerlinde) Haverkamp (2017). Depressive symptoms in an ethnically DIVERSe cohort of chronic dialysis patients: The role of patient characteristics, cultural and inflammatory factors. Vrije Universiteit Amsterdam. ISBN: 978-94-6233-528-8.

T.J. (Tjalling) Holwerda (2017). Burden of loneliness and depression in late life. Vrije Universiteit Amsterdam. ISBN: 978-94-6233-598-1.

J. (Judith) Verduijn (2017). Staging of Major Depressive Disorder. Vrije Universiteit Amsterdam. ISBN: 978-94-6299-597-0.

C.N. (Catherine) Black (2017). Oxidative stress in depression and anxiety disorders. Vrije Universiteit Amsterdam. ISBN: 978-94-6299-672-4.

J.B. (Joost) Sanders (2017). Slowing and Depressive Symptoms in Aging People. Vrije Universiteit Amsterdam. ISBN: 978-94-6233-650-6.

W. (Willemijn) Scholten (2017). Waxing and waning of anxiety disorders: relapse and relapse prevention. Vrije Universiteit Amsterdam. ISBN: 978-94-6299-606-9.

P. (Petra) Boersma (2017). Person-centred communication with people with dementia living in nursing homes; a study into implementation success and influencing factors. Vrije Universiteit Amsterdam. ISBN: 978-94-6233-725-1.

T.I. (Annet) Bron (2017). Lifestyle in adult ADHD from a Picasso point of view. Vrije Universiteit Amsterdam. ISBN: 978-94-6299-685-4.

S.W.N. (Suzan) Vogel (2017). ADHD IN ADULTS: seasons, stress, sleep and societal impact. Vrije Universiteit Amsterdam. ISBN: 978-94-6299-673-1.

R. (Roxanne) Schaakxs (2018). Major depressive disorder across the life span: the role of chronological and biological age. Vrije Universiteit Amsterdam. ISBN: 978-94-6299-819-3.

J.J. (Bart) Hattink (2018). Needs-based enabling- and care technology for people with dementia and their carers. Vrije Universiteit Amsterdam. ISBN: 978-94-6295-880-7.

F.T. (Flora) Gossink (2018). Late Onset Behavioral Changes differentiating between bvFTD and psychiatric disorders in clinical practice. Vrije Universiteit Amsterdam. ISBN: 978-94-6295-899-9.

R. (Roxanne) Gaspersz (2018). Heterogeneity of Major Depressive Disorder. The role of anxious distress. Vrije Universiteit Amsterdam. ISBN: 978-94-028-1076-9.

M.M. (Marleen) Wildschut (2018). Survivors of early childhood trauma and emotional neglect: who are they and what's their diagnosis? Vrije Universiteit Amsterdam. ISBN: 978-94-6332-401-4.

J.A.C. (Jolanda) Meeuwissen (2018). The case for stepped care. Exploring the applicability and cost-utility of stepped-care strategies in the management of depression. Vrije Universiteit Amsterdam. ISBN: 978-90-5383-359-9.

D.S. (Dora) Wynchank (2018). The rhythm of adult ADHD. On the relationship between ADHD, sleep and aging. Vrije Universiteit Amsterdam. ISBN: 978-94-6375-034-9.

M.J. (Margot) Metz (2018). Shared Decision Making in mental health care: the added value for patients and clinicians. Vrije Universiteit Amsterdam. ISBN: 978-94-6332-403-8.

I. (Ilse) Wielaard (2018). Childhood abuse and late life depression. Vrije Universiteit Amsterdam. ISBN: 978-94-6380-072-3.

L.S. (Laura) van Velzen (2019). The stressed and depressed brain. Vrije Universiteit Amsterdam. ISBN: 978-94-6380-062-4.

S. (Sonja) Rutten (2019). Shedding light on depressive, anxiety and sleep disorders in Parkinson's disease. Vrije Universiteit Amsterdam. ISBN: 978-94-6380-176-8.

N.P.G. (Nadine) Paans (2019). When you carry the weight of the world not only on your shoulders. Factors associating depression and obesity. Vrije Universiteit Amsterdam. ISBN: 978-94-6380-141-6.

D.J. (Deborah) Gibson-Smith (2019). The Weight of Depression. Epidemiological studies into obesity, dietary intake and mental health. Vrije Universiteit Amsterdam. ISBN: 978-94-6380-144-7.

C.S.E.W. (Claudia) Schuurhuizen (2019). Optimizing psychosocial support and symptom management for patients with advanced cancer. Vrije Universiteit Amsterdam. ISBN: 978-94-6323-468-9.

M.X. (Mandy) Hu (2019). Cardiac autonomic activity in depression and anxiety: heartfelt afflictions of the mind. Vrije Universiteit Amsterdam. ISBN: 978-94-6380-206-2.

J..K. (Jan) Mokkenstorm (2019). On the road to zero suicides: Implementation studies. Vrije Universiteit Amsterdam. ISBN: 978-94-6361-224-1.

S.Y. (Sascha) Struijs (2019). Psychological vulnerability in depressive and anxiety disorders. Vrije Universiteit Amsterdam. ISBN: 978-94-6380-244-4.

H.W. (Hans) Jeuring (2019). Time trends and long-term outcome of late-life depression: an epidemiological perspective. Vrije Universiteit Amsterdam. ISBN: 978-94-6380-228-4.

R. (Ruth) Klaming Miller (2019). Vulnerability of memory function and the hippocampus: Risk and protective factors from neuropsychological and neuroimaging perspectives. Vrije Universiteit Amsterdam. ISBN: 978-94-6182-955-5.

P.S.W. (Premika) Boedhoe (2019) The structure of the obsessive-compulsive brain – a worldwide effort. Vrije Universiteit Amsterdam. ISBN: 978-94-6380-329-8.

C.S. (Carisha) Thesing (2020). Fatty acids in depressive and anxiety disorders: fishing for answers. Vrije Universiteit Amsterdam. ISBN: 978-94-6375-846-8.

R.D. (Richard) Dinga (2020). Evaluation of machine learning models in psychiatry. Vrije Universiteit Amsterdam.

M. (Mayke) Mol (2020). Uptake of internet-based therapy for depression: the role of the therapist. Vrije Universiteit Amsterdam. ISBN: 978-94-6416-150-2.

R.C. (Renske) Bosman (2020). Improving the long-term prognosis of anxiety disorders: Clinical course, chronicity and antidepressant use. Vrije Universiteit Amsterdam. ISBN: 978-94-6375-736-2.

R.W. (Robbert) Schouten (2020). Anxiety, depression and adverse clinical outcomes in dialysis patients. Should we do more? Vrije Universiteit Amsterdam. ISBN: 978-94-6416-179-3.

T.T. (Trees) Juurlink (2021). Occupational functioning in personality disorders: a quantitative, qualitative and semi-experimental approach. Vrije Universiteit Amsterdam. ISBN: 978-94-6421-117-1.

I.P.H. (Ires) Ghielen (2021). Surfing the waves of Parkinson's disease. Understanding and treating anxiety in the context of motor symptoms. Vrije Universiteit Amsterdam. ISBN: 978-94-6416-493-0.

L.K.M. (Laura) Han (2021). Biological aging in major depressive disorder. Vrije Universiteit Amsterdam. ISBN: 978-94-93184-91-6.

E. (Esther) Krijnen-de Bruin (2021). Relapse prevention in patients with anxiety or depressive disorders. Vrije Universiteit Amsterdam. ISBN: 978-94-6423-298-1

T.D. (Tim) van Balkom (2021). The profiles and practice of cognitive function in Parkinson's disease. Vrije Universiteit van Amsterdam. ISBN: 978-94-6423-391-9

S.M. (Sanne) Swart (2021). The course of survivors of early childhood trauma and emotional neglect: never easy, but worth it? Vrije Universiteit Amsterdam. ISBN: 978-94-6416-650-7

Y.J.F. (Yvonne) Kerkhof (2021). Digital support for self-management and meaningful activities of people with mild dementia. Development, implementation and feasibility of a person-centred touch-screen intervention. Vrije Universiteit Amsterdam. ISBN: 978-90-829978-2-8

I.M.J. (Ilja) Saris (2021). Together alone: Social dysfunction in neuropsychiatric disorders. Vrije Universiteit Amsterdam. ISBN: 978-90-9035-072-1

A. (Angela) Carlier (2021). Biomarkers and electroconvulsive therapy in late-life depression. Vrije Universiteit Amsterdam. ISBN: 978-94-6421-462-8

S. (Sonia) Difrancesco (2021). Sleep, circadian rhythms and physical activity in depression and anxiety.The role of ambulatory assessment tools. Vrije Universiteit Amsterdam. ISBN: 978-94-6416-781-8

B.A. (Bianca) Lever-van Milligen (2021). The interplay between depression, anxiety and objectively measured physical function. Vrije Universiteit Amsterdam. ISBN: 978-94-6423-443-5

J.M. (Joeke) van der Molen-van Santen (2021). Remember to play... and stay active! Evaluation of the effects, cost-effectiveness and implementation of exergaming for people living with dementia and their informal caregivers. Vrije Universiteit Amsterdam. ISBN: 978-94-6332-795-4

W.A. (Wicher) Bokma (2021). Worrying about the future: towards evidence-based prognosis in anxiety disorders. Vrije Universiteit Amsterdam. ISBN: 978-94-93270-17-6