Staging and profiling in clinical practice:

To what extent do items on a checklist of staging and profiling predict treatment course in anxiety patients? An exploratory study.

Claire Kortenoeven-Chillaud, BSc

July 24th 2015

Objective: Up to today we remain a long way from identifying the most robust predictors of drop-out and (non-)response for treatment. To enable assigning patients to the right intensity of treatment, the Altrecht Academic Anxiety center (AAA) has developed a Checklist Staging and Profiling (CSP), based on consensus between clinicians and common sense. This exploratory study has investigated whether the CSP can differentiate between different disease stages in anxiety disorders and predict treatment outcome in anxiety patients, based on their stage of illness. **Methods**: The current study consisted of a retrospective study – using the various questionnaires on symptoms at baseline and follow-up of the AAA – and a prospective study of two months duration. Both studies included anxiety patients of the AAA (specialized mental health care), whose data on background variables and symptoms were collected. **Results**: The CSP has sufficient internal consistency, and differentiates accurately between the different stages of anxiety disorders. The CSP can reasonably predict the course of treatment in patients with anxiety disorders and also which patients will drop-out if no appropriate measures are taken. **Discussion**: With some adaptations, the CSP has a lot of potential. In future research, the improved version in Appendix F should be tested in clinical practice to further enhance the usefulness of the CSP. Further, we recommend to add items on patients' belief in the treatment rationale, motivation for treatment and social support to the CSP and study whether this would enhance its predictability.

Doelstelling: Tot op de dag van vandaag hebben we nog een lange weg te gaan voordat we de meest robuuste voorspellers kunnen identificeren van uitval en therapieresistentie. Om het mogelijk te maken patiënten de juiste intensiteit van behandeling toe te wijzen, heeft het Altrecht Academisch Angstcentrum (AAA) de Checklist Stagering en Profilering (CSP) ontwikkelt, gebaseerd op consensus tussen clinici en algemene kennis. Deze explorerende studie heeft onderzocht of de Checklist Stagering en Profilering (CSP) kan differentiëren tussen de verschillende ziektestadia in angststoornissen en behandeluitkomst bij angstpatiënten kan voorspellen, gebaseerd op hun ziektestadium. **Methoden**: Het huidige onderzoek bestond uit een retrospectief onderzoek – gebruikmakend van het elektronisch patiëntendossier van het AAA – en een prospectief onderzoek. Beide onderzoeken maakten gebruik van angstpatiënten binnen Altrecht (specialistische GGz), wiens data betreffende achtergrondvariabelen en symptomen waren verzameld. Er werden meerdere vragenlijsten ingevuld op verschillende momenten. **Resultaten**: De CSP heeft voldoende interne consistentie, en differentieert accuraat tussen de verschillende stadia van angststoornissen. De CSP kan de behandeluitkomst voorspellen in patiënten met angststoornissen, en ook welke patiënten zullen uitvallen indien er geen passende maatregelen worden genomen. **Discussie**: Met enkele aanpassingen heeft de CSP veel potentie. In toekomstig onderzoek zou de verbeterde versie in Appendix F getest moeten worden in de klinische praktijk om zo de bruikbaarheid van de CSP te verbeteren. Daarnaast raden wij aan om items over de overtuigingen van patiënten over de rationale van de behandeling, de behandelmotivatie en sociale steun toe te voegen aan de CSP en te onderzoeken of dit toegevoegde, voorspellende waarde heeft.

Acknowledgements

I would like to thank both dr. Tonnie Staring and dr. Daniëlle Cath for their support and guidance throughout the completion of my master research and thesis. It was quite an adventure, full of obstacles, but in the end we made it. I would also like to thank Naomi Paauw for our collaboration throughout this process. Particularly for helping me to collect the data. Further, I would like to thank all members of the methodology & statistics department who assisted me throughout this process, and also a special thanks to dr. Shahab Jolani who helped me with the data imputation I'd never learned before and I couldn't have done without him.

Index

Anxiety disorders are common, often disabling, and costly in terms of personal suffering and demands on the health care system (Tolin, Gilliam & Dufresne, 2010). With an average prevalence of 10.8% in the total population, anxiety disorders are the most common disorders amongst adults according to a cross-sectional study (2001-2009) of a representative sample of adults in 24 countries worldwide of the World Health Organization (Wang, et al., 2011). According to NEMESIS-2 (Netherlands Mental Health Survey and Incidence Study-2) of the Trimbos Institute (a national Dutch knowledge Institute of mental health care) the lifetime prevalence of anxiety disorders in The Netherlands is 19.6% and the one-year incidence was 10.1% in 2009 (De Graaf, Ten Have & Van Dorsselaer, 2010). A meta-analysis of 41 prevalence and 5 incidence studies performed by Somers, Goldner, Waraich and Hsu (2006) showed similar figures, with pooled 1-year and lifetime prevalence rates for total anxiety disorders of 10.6% and 16.6% respectively. Although an insufficient number of incidence studies were available to clarify details concerning the onset of anxiety symptoms worldwide, Kessler, Ruscio, Shear and Wittchen (2009) showed in their review that anxiety disorders typically begin early in life. Most specific phobias have their onset in childhood, with the vast majority of lifetime cases having onsets by the age of 18. Social phobia and obsessive-compulsive disorder have their onset in adolescence or early adulthood, with the vast majority of lifetime cases beginning by their twenties. Panic disorder, agoraphobia and generalized anxiety disorder have a more dispersed onset distribution, with most of lifetime cases beginning in their twenties. Posttraumatic stress disorder is generally found to have the latest and most variable onset distribution, presumably reflecting the fact that trauma exposure can occur at any time in the course of life. It is noteworthy that despite their generally early age of onset, first treatment of anxiety disorders usually does not occur until adulthood, often more than a decade after the onset of the disorder (Christiana et al. 2000).

Anxiety disorders are defined by abnormal fear, in which the fear gives rise to sustained subjective distress or impedes social functioning. There are several types of anxiety disorders. The current study follows the internationally accepted classification of anxiety disorders according to the DSM-IV-TR, which differentiates panic disorder (sudden and repeated feelings of terror with physical symptoms like sweating and heartbeats), agoraphobia (fear of any situation where escaping is difficult, along with avoidance and safety behaviors), specific phobia (intense and unreasonable fear of a specific object or situation), social anxiety disorder (persistent overwhelming worry and self-consciousness about everyday social

situations), obsessive-compulsive disorder (intrusive thoughts that cause unreasonable fears and repetitive behaviors aimed at reducing the associated anxiety), posttraumatic stress disorder (recurring, frightening thoughts or memories of a traumatic event, including emotionally numbness, nightmares and physiological fear reactions) and generalized anxiety disorder (excessive, unrealistic worry and tension, even if there is nothing to provoke the anxiety, that causes physical stress symptoms; American Psychiatric Association, 2001).

The ESEMeD study (European Study of the Epidemiology of Mental Disorders), a cross-sectional study (2001-2003) of a representative sample of adults in The Netherlands, Belgium, Germany, France, Italy and Spain, shows that the overall proportion of adequate treatment for anxiety disorders in Europe was on average 54.5% for the general mental care and specialized mental care sectors together. Treatment adequacy for both anxiety and depressive disorders in the general mental care sector was on average 23.3% and in the specialized mental care sector 57.4% (criteria for minimally adequate treatment were: receiving antidepressant or anxiolytic pharmacotherapy for at least two months and at least four visits with a psychiatrist, general practitioner or other doctor; or at least eight sessions with a psychologist or psychiatrist lasting an average of thirty minutes), which shows that patients often don't get the help they need (Fernández, et al., 2007). Another study shows the same results for The Netherlands: in moderate to severe anxiety disorders, 44% of the subjects were not treated sufficiently with medication or psychological treatment. Treatment inadequacy was more prominent in general health care (60%) than in specialized health care (30%; Bet, et al., 2013). The result is that they stay in treatment much longer than expected, they drop-out halfway or relapse afterwards. Several studies reported dropout rates from different types of anxiety treatment up to 88%, with the vast majority being in the 15-30% drop-out range (Høifødt, Strøm, Kolstrup, Eisemann & Waterloo, 2011; Santana & Fontenelle, 2011). Meta-analyses of anxiety disorders have reported dropout rates in the range of 9–21% for CBT and 18–30% for SRIs (Taylor, Abramowitzb & McKay, 2012). One of the factors that are assumed to contribute to this lack of treatment success in patients is that patients are not adequately allocated to the treatment modalities and intensity they need. Can we decrease drop-out and relapse rates by allocating patients to the adequate treatment intensity at an early stage of treatment? And can we predict which patients will react well on an anxiety treatment and which patients will dropout, and subsequently design alternatives for the potential drop-outs at an early phase of treatment allocation?

One of the assumptions we make is that the longer and more severely affected patients are, the more intensive the treatment they need. We expect that the stage of the anxiety disorder – that includes severity, duration and the like – and the profile of the patient's characteristics – including age, GAF score and the like – at the start of treatment can predict the course of treatment. Van Balkom and colleagues (2012) were the first in the Netherlands to propose a staging model for patients with anxiety disorders. In this article we will reserve the term 'staging' to classify differences in severity of anxiety disorders into different stages. Figure 1 shows this model, which is subdivided into four stages, either with favorable or unfavorable courses. The first stage is the least severe stage and the fourth stage is the most severe one.

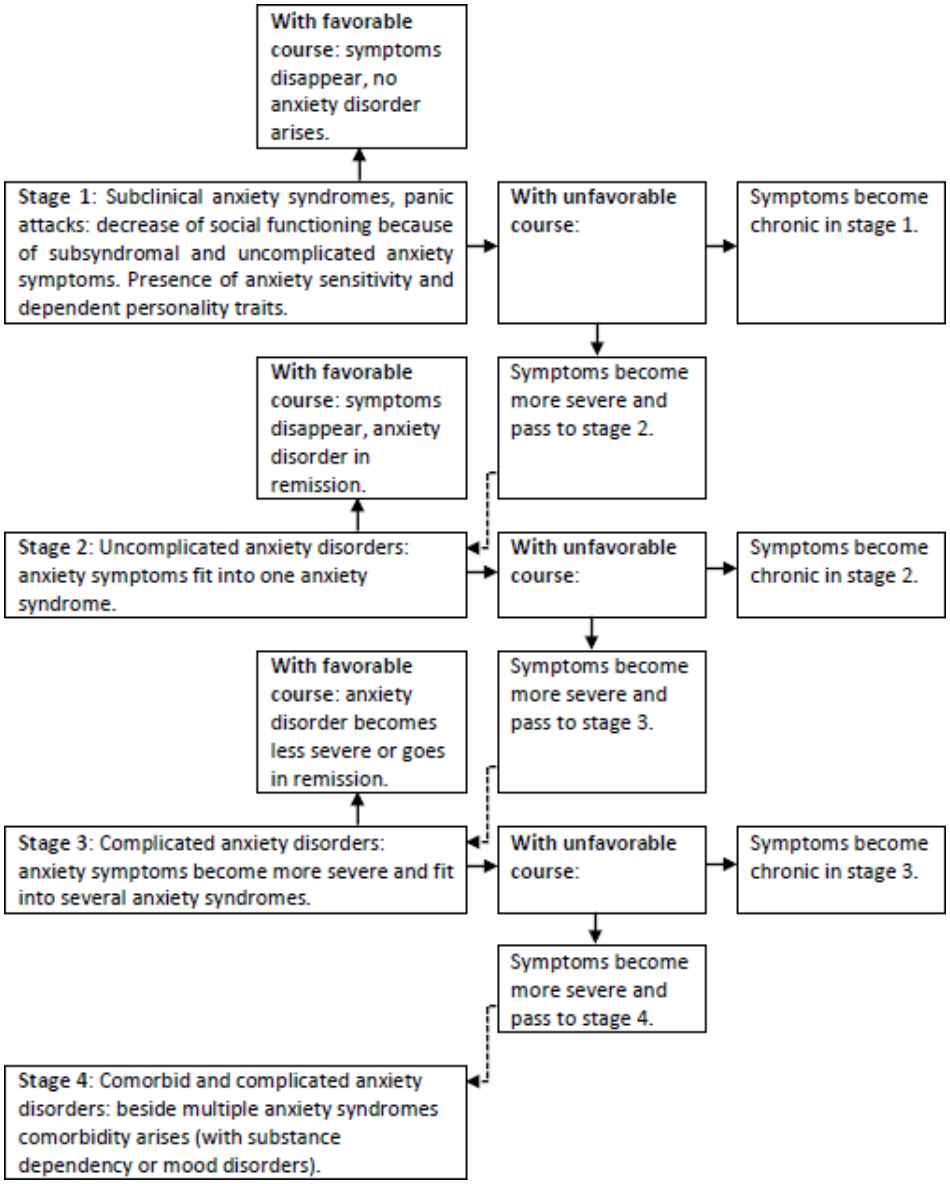

Figure 1. The staging model of Van Balkom and colleagues (2012)

Staging is important for clinical practice; when we can predict what patients need, how they will respond to treatment and which patients will profit from more intensive as opposed to less intensive treatment, we can reduce treatment duration and drop-out. This would not only be great for the patients themselves, but also for the costs of mental health care. Van Balkom and colleagues (2012) explain that clinicians generally base their decisions for treatment on the individual characteristics of the patient and his anxiety disorder, but they don't do it based on scientific insights. The Dutch guideline for Anxiety Disorders, existing since 2003, gives treatment guidelines based on severity of the disorder, on co-morbid depression and on the effect of previous treatments, without a proper scientific foundation as well (Van Balkom, et al., 2013). Checklists on staging are currently not available, but staging of patients is important to decide on intensity of treatment and, accordingly, to decide whether a patient should be treated in basic or in specialized mental health care. Therefore, based on a consensus meeting by clinicians from the Altrecht Academic Anxiety (AAA) outpatient clinics a checklist was designed: the Checklist Staging and Profiling (CSP[1]). This checklist has the aim to decide in a more rational way on whether patients should receive specialized treatment, and on treatment intensity. As the AAA encompasses a specialized treatment center aiming to treat treatment resistant patients, the assumption at our department was that such patients can mostly be treated adequately, on the premise that this encompasses an intensive and specialized treatment that targets their symptoms rather than presumed underlying processes. However, we are in need of an instrument that is able to assess and detect those factors that contribute to assigning to the adequate treatment and predict whether a treatment is successful. If the CSP is able to accurately determine the stage of an anxiety disorder, patients will get the treatment they need and recover sooner from their illness. But how can we predict which patients will need a more intensive treatment? Which factors affect the course of treatment?

The CSP contains 13 items that include variables that are judged by practitioners to be clinically relevant for predicting the course of treatment (Table 1, next page). Further, most items are supported by research findings on treatment refractoriness. To summarize, items 2, 3 and 11 reflect number of co-morbid conditions (A), items 4, 5, 6 and 9 reflect illness severity (B), item 7 age (C), item 8 illness duration (D), items 10a and 10c number and effect of previous treatments (E) and items 12 and 13 level of premorbid functioning, including

[1] See Table 1 and Appendix A for the items of the CSP.

level of education (F). Illness severity (B) and the number of symptoms, including comorbid disorders (A), have proven to be predictors of non-response to treatment (Slaap, Van Vliet, Westenberg & Den Boer, 1996; Taylor, Abramowitz & McKay, 2012; Mululo, Bezerra de Menezes, Paula & Fontenelle, 2012). White and colleagues (2010) found that older age (C) predicted higher odds of completing treatment in panic disorder. With respect to illness duration, it has been found that a duration (D) of an untreated panic disorder longer than one year was linked to more frequent comorbidity with major depressive disorder (Altamura, Santini, Salvadori and Mundo, 2005), reduced response to antidepressants and worse long-term outcome (Altamura, et al., 2008; as cited in Altamura, Buoli, Albano & Dell'Osso, 2010).

Table 1

Items of the Checklist Staging and Profiling (CSP)

Items	**Score**
1) Are there multiple anxiety disorders?	0-1
2) Is there a comorbid depression / dysthymia?	0 or 2
3) What is the severity of the (main) anxiety disorder?	0-2
4) In how many life areas does the (main) disorder affects daily functioning?	0-3
5) In how may situation does the patient use avoidance behavior?	0-3
6) Age	0-4
7) Duration of symptoms or complaints	0-4
8) Course of the (main) disorder: fluctuating or chronic?	0-3
9) How many previous treatments (of every kind)?	0-2
10) Effect of previous treatment (considered as a whole)?	0-2
11) Are there other comorbidities (beside anxiety disorders and depression)?	0-3
12) GAF score at premorbid functioning or intake	0-3
13) Highest education completed	0-3

The number of previous treatments and their effect (E) also seem clinically important, but have not been examined as a relevant treatment outcome predictor before (Van Balkom, et al., 2012). Furthermore, there are some conflicting findings about the role of education (F) as a predictor for treatment outcome. For example, McLean and colleagues (2001) found that treatment completers were relatively well educated: most had at least some postsecondary

education. However, Ebert and colleagues (2013) found that treatment is especially effective for participants with low education levels. So it's interesting to investigate if education level is a predictor for treatment outcome. Smith, Van Ryzin, Fowler and Handler (2014) found that patients with lower premorbid Global Assessment Functioning (GAF) scores showed little to no improvement. To conclude, this literature mostly supports the relevance of the variables included in the CSP, to predict the stage of illness that the patient is in. Whether the clinicians-constructed CSP meets these goals forms the subject of the current thesis.

We aimed to investigate the relative contribution of the items of the CSP to the correct prediction of the illness stage of a patient's anxiety disorder. Prior to investigating this, we studied the association between the categories of the CSP and illness severity and treatment refractoriness. This way, we checked the validity of these categories to indicate stage of the illness, that is: uncomplicated v/s complicated disorder and a favorable v/s unfavorable prognosis. It was expected that (1) the CSP predicts the course of treatment in patients with anxiety disorders, and that (2) the CSP also predicts drop-out in patients with anxiety disorders, in the following way: patients with high scores on the CSP (e.g. the third and fourth category) would be less likely to make progress in their treatment and be more likely to drop out; the first category would contain the patients with the least illness severity and the fourth category would contain patients with the highest illness severity.

Methods

Design

The current study consisted of a retrospective (retro) and a prospective study (pro). An attempt was made to replicate the results from the retrospective study in the prospective study. In both studies the predictive value of the CSP for treatment outcome, that is reduction in anxiety, was studied. In the retrospective as well as the prospective study the score on the CSP was the independent variable, both the continuous (0-35 points) and the categorical variable (four categories), and the scores on the anxiety measures were the dependent variables. The dependent variables consisted of six (sub)scales: OQ45 total score, BSI total score, BSI depressive subscale, BSI anxiety subscale, BSI phobic anxiety subscale and BDI total score. The CSP was measured only once for each participant, in the retrospective study based on the electronic patient dossier information at baseline, and in the prospective study at the beginning of their therapy. The dependent variables were measured at least two times for each participant, comparing baseline-measurement (T_0) with a measurement after two months treatment (T_1) in the prospective study, and also with a post-treatment measurement (T_2) and follow-up measurement (T_3) in the retrospective study.

Material

Both studies included the following questionnaires:

Checklist Staging and Profiling

The Checklist Staging and Profiling (CSP; dr. D. Cath, head of the AAA, personal correspondence, September 24th 2013) entailed an interview supposed to be taken at the start of the treatment from the treating clinicians and consisted of 13 items about the patient and his anxiety disorder (see Table 1 in preface) and the total score ranged from 0 to 35. Scores on the CSP yield four categories which are somewhat similar to the stages proposed by Van Balkom and colleagues (2012): uncomplicated anxiety disorder with favorable prognosis (0-14 points, first category), uncomplicated anxiety disorder with unfavorable prognosis (15-22 points, second category), complicated anxiety disorder with favorable prognosis (23-28 points, third category) and complicated anxiety disorder with unfavorable prognosis (28-35 points, fourth category). Patients in the first category are considered to be adequately treated within the context of general mental health care, those in the second category in specialized

(secondary) mental health care, the third category is regarded as needing a more intensive outpatient treatment in specialized (secondary or tertiary/topclinical) mental health care and the fourth category should get a highly intensive treatment or support according to a handicap model.

In the prospective study, the master students (C.K. and N.P.) interviewed the professionals (mostly psychologists or cognitive behavioral therapists) involved in the treatment of patients, at the beginning of treatment, with the aid of the CSP. In the retrospective study, all the answers to the CSP items were derived from the electronic health records of the AAA. The item about the severity of the anxiety disorder (item 4) was derived from the total score of the Anxiety Sensitivity Index (ASI-3; Taylor et al., 2007), derived at baseline.

Anxiety Sensitivity Index

This questionnaire was solely used to calculate its correlation with the CSP (for missing data and validity, with all scales) and to calculate the CSP item about the severity of the anxiety disorder (using the total scale). For the item about the severity of the anxiety disorder (item 4) the total score at baseline was used as followed: 0-15 points for a mild (0), 16-30 points for a moderate (1) and 31-72 points for a severe (2) anxiety disorder. The Anxiety Sensitivity Index-3 (ASI-3; Taylor et al., 2007) is an 18-item version of the original ASI (Reiss et al., 1986) that measures beliefs about the feared consequences of symptoms associated with anxious arousal. Patients indicated their agreement with each item from "very little" (coded as 0) to "very much" (coded as 4). Total scores range from 0 to 72. The measure possesses excellent psychometric properties, performing well on various indices of reliability and validity (Taylor et al., 2007). Reliability for the total score is excellent (R^2 = .93; Nunnally & Bernstein, 1994).

Outcome Questionnaire 45

The Outcome Questionnaire 45 (OQ45; Wells, Burlingame, Lambert, Hoag & Hope, 1996) consists of 45 items that measures general emotional and lifestyle stressors (such as depression, anxiety, stress, substance abuse and suicidality), satisfaction with relationships (marital, family and friendships), work relations and leisure activities over the past week. Patients indicated the frequency of each item from "never" (coded as 0) to "almost always" (coded as 4). Total scores range from 0 to 180. The OQ45 contains three empirically

established subscales assessing symptom distress (e.g., "I feel anxious"), interpersonal relations (e.g., "I feel lonely"), and social role (e.g., "I am under stress at work/school"). Scores on the complete OQ45 scale have been reported to be reliable and valid, distinguishing well between clinical and non-clinical subjects (Umphress, Lambert, Smart & Barlow, 1997). Internal consistency and test-retest reliability estimates range from .70 to .93 and .78 to .84, respectively (Wells et al., 1996), showing good to excellent psychometric properties (Nunnally & Bernstein, 1994).

Brief Symptom Inventory

The Brief Symptom Inventory (BSI; Derogatis, 1975) consists of 53 items that measures to what extent the patient was suffering of mental and/or physical symptoms over the past period of time (e.g., "During the past week, how much did you suffer from nervousness or shakiness?"). Patients indicated the frequency of each item from "never" (coded as 0) to "very much" (coded as 4). Total scores range from 0 to 212. The BSI contains nine empirically established subscales assessing somatic complaints, cognitive problems, interpersonal sensitivity, depressive mood, anxiety, hostility, phobic anxiety, paranoid thoughts, and psychoticism. The BSI has been reported to be sufficiently reliable and valid (De Beurs & Zitman, 2006).

Beck Depression Inventory

The Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) consists of 21 items that measures the severity of a depressive disorder over the past two weeks (e.g., "I don't feel gloomy / I feel gloomy a lot of times / I feel gloomy all the time / I feel so gloomy or unhappy that I can't bare it anymore"). Patients indicated the severity of each item from "never" (coded as 0) to "very much" (coded as 3). Total scores range from 0 to 63. Studies reviewed by Steer, Beck, and Garrison (1986) have supported internal consistencies in the .90s and high positive correlations with clinical ratings of depression, showing excellent psychometric properties (Nunnally & Bernstein, 1994).

Complementary questionnaire used in the prospective study only:

Beck Anxiety Inventory

This questionnaire was solely used to calculate it's correlation with the CSP (with all scales). The Beck Anxiety Inventory (BAI; Beck, Epstein, Brown & Steer, 1988) consists of 21 items

that measures the severity of, especially somatic symptoms of, anxiety. Patients indicated their agreement with each item from "not at all" (coded as 0) to "severely" (coded as 3). Total scores range from 0 to 63. The BAI contains two empirically established subscales relating to the somatic symptoms of anxiety (e.g., "Dizzy or light-headed") and the subjective characteristics of anxiety (e.g., "Afraid that the worst happens"). The BAI scale has a high internal consistency (α = .92) and item-total correlations ranging from .30 to .71. The test-retest reliability for the total score is .75, showing good psychometric properties (Beck et al., 1988; Nunnally & Bernstein, 1994).

Participants

The retrospective study

All participants were patients of the AAA. One hundred and fifty patients were randomly selected from a total of 737 patients, based on the following criteria: patients had had treatment between 2008 and 2013, were treated for an anxiety disorder and had had at least one measurement, at baseline. The measurements at baseline entailed multiple psychodiagnostic measurements to evaluate their symptoms. Of these 150 patients 45 patients just had had the baseline measurement, 50 patients had had two measurements, 49 patients had had three measurements and 6 patients had had all four measurements. Two-thirds of the patients that have taken part in this study were women (68.0%), mean age was 33 years, and men and women did not differ in age (Table 2).

Table 2

Number, Age and Country of Origin broken down by Gender (Retrospective Study)

	Men	Women	Total
N	48	102	150
Min. – max. age	22 – 56	18 – 57	18 – 57
Age *M* (*SD*)	35.17 (9.78)	32.08 (9.26)	33.07 (9.50)
Native	38 (79.2)	70 (68.6)	108 (72.0)
First generation immigrant	8 (16.7)	18 (17.6)	26 (17.3)
Second generation immigrant	2 (4.2)	14 (13.7)	16 (10.7)
The Netherlands	40 (83.3)	82 (80.4)	122 (81.3)
Turkey	0 (0.0)	4 (3.9)	4 (2.7)
Morocco	3 (6.3)	10 (9.8)	13 (8.7)
Surinam	1 (2.1)	1 (1.0)	2 (1.3)
The Dutch Antilles	0 (0.0)	1 (1.0)	1 (0.7)
Other country	4 (8.3)	4 (3.9)	8 (5.3)

The vast majority (ca. 70-80%) of patients was Dutch or had The Netherlands as their country of origin (both parents being Dutch as well).

Fifty-three patients (35.3%) scored in the first category, 65 patients (43.3%) in the second category, 28 (18.7%) in the third category and 4 patients (2.7%) in the fourth category of the CSP. Most patients had only one anxiety disorder (56.7%) and were suffering from panic disorder with agoraphobia (26.0%), social anxiety disorder (23.3%) or obsessive-compulsive disorder (22.0%, Table 3). In those with comorbid anxiety disorders specific phobia (12.7%), generalized anxiety disorder (11.3%) and social anxiety disorder (10.0%) were the most common additional anxiety disorders (see Appendix C, Table 16). Sixty percent of patients had at least one non-anxiety comorbid disorder (see Appendix C, Table 18), major depressive disorder or dysthymia being the most common additional comorbid disorder (46.0%, Table 3). Almost 75% of patients were working and or studying, and 42% of patients was living alone. See Table 4 for further details with respect to work and family status.

Table 3

CSP Category, Main Diagnosis, Comorbid Anxiety and Depressive Disorders broken down by Gender (Retrospective Study)

	Men *N* (%)	Women *N* (%)	Total *N* (%)
First CSP category	18 (37.5)	35 (34.3)	53 (35.3)
Second CSP category	17 (35.4)	48 (47.1)	65 (43.3)
Third CSP category	13 (27.1)	15 (14.7)	28 (18.7)
Fourth CSP category	0 (0.0)	4 (3.9)	4 (2.7)
Panic disorder (without agoraphobia)	2 (4.2)	0 (0.0)	2 (1.3)
Panic disorder with agoraphobia	14 (29.2)	25 (24.5)	39 (26.0)
Social anxiety disorder (social phobia)	12 (25.0)	23 (22.5)	35 (23.3)
Obsessive-compulsive disorder	12 (25.0)	21 (20.6)	33 (22.0)
Posttraumatic stress disorder	3 (6.3)	17 (16.7)	20 (13.3)
Generalized anxiety disorder	3 (6.3)	12 (11.8)	15 (10.0)
Anxiety disorder NAO	2 (4.2)	4 (3.9)	6 (4.0)
One anxiety disorder	36 (75.0)	49 (48.0)	85 (56.7)
Two anxiety disorders	10 (20.8)	40 (39.2)	50 (33.3)
Three anxiety disorders	2 (4.2)	12 (11.8)	14 (9.3)
Four anxiety disorders	0 (0.0)	1 (1.0)	1 (0.7)
Major depressive disorder or dysthymia	21 (43.8)	48 (47.1)	69 (46.0)

Table 4

Work and Family Status broken down by Gender (Retrospective Study)

	Men *N* (%)	Women *N* (%)	Total *N* (%)
Working	35 (72.9)	52 (51.0)	87 (58.0)
Studying	2 (4.2)	16 (15.7)	18 (12.0)
Working and studying	3 (6.2)	4 (3.9)	7 (4.7)
Neither working nor studying	8 (16.7)	30 (29.4)	38 (25.3)
Living alone	14 (29.2)	28 (27.5)	42 (28.0)
Living with partner	14 (29.2)	26 (25.5)	40 (26.7)
Living (alone) with children	0 (0.0)	12 (11.8)	12 (8.0)
Living with partner and children	14 (29.2)	19 (18.6)	33 (22.0)
Living with parents/guardians	3 (6.3)	10 (9.8)	13 (8.7)
Living with other person	3 (6.3)	7 (6.9)	10 (6.7)

The prospective study

Thirty three patients of the AAA were recruited for the prospective study of whom 12 were excluded because of the requirements: five have had more than five sessions at the time of completing the CSP, three had not had a baseline measurement and four had had a baseline measurement after the start of their treatment. In the end, 21 patients participated in the study, of whom two dropped out during treatment (and did not participate at T_1). Conclusively, 19 patients stayed in treatment during the course of the study, with an average age of 39 years, containing 9 men and 12 women (table 5). The vast majority (ca. 75-85%) of patients was Dutch or had The Netherlands as their country of origin.

Table 5

Number, Age and Country of Origin broken down by Gender (Prospective Study)

	Men	Women	Total
N	9	12	21
Minimum – maximum age	28 – 60	20 – 57	20 – 60
Age *M* (*SD*)	45.44 (11.80)	34.00 (10.03)	38.90 (12.03)
Native	--	--	16 (76.2)
First generation immigrant	--	--	2 (9.5)
Second generation immigrant	--	--	3 (14.3)
The Netherlands	--	--	18 (85.7)
Morocco	--	--	2 (9.5)
Other country	--	--	1 (4.8)

One patient (4.8%) scored in the first category, 8 patients (38.1%) in the second category, 11 (52.4%) in the third category and 1 patient (4.8%) in the fourth category of the CSP at baseline. Most patients had only one anxiety disorder (66.7%) and were suffering from social anxiety disorder (33.3%) or posttraumatic stress disorder (28.6%, Table 6). In those with comorbid anxiety disorders obsessive-compulsive disorder (14.3%) and specific phobia (9.5%) were the most common additional anxiety disorders (see Appendix C, Table 17). Eighty-five percent of patients had at least one non-anxiety comorbid disorder (see Appendix C, Table 19). Of those, major depressive disorder or dysthymia was the most common additional comorbid disorder (42.9%, Table 6). See Table 6 for further details with respect to work and family status.

Table 6

Main Diagnosis, Comorbid Anxiety and Depressive Disorders, Work and Family Status (Prospective Study)

	Total *N* (%)
Panic disorder with agoraphobia	1 (4.8)
Social anxiety disorder (social phobia)	7 (33.3)
Obsessive-compulsive disorder	4 (19.0)
Posttraumatic stress disorder	6 (28.6)
Generalized anxiety disorder	2 (9.5)
Anxiety disorder NAO	1 (4.8)
One anxiety disorder	14 (66.7)
Two anxiety disorders	6 (28.6)
Three anxiety disorders	1 (4.8)
Major depressive disorder or dysthymia	9 (42.9)
Working	10 (47.6)
Studying	4 (19.0)
Neither working nor studying	7 (33.3)
Living alone	2 (9.5)
Living with partner	9 (42.9)
Living (alone) with children	4 (19.0)
Living with partner and children	6 (28.6)

Procedure

The retrospective study

As part of the standard procedure in their treatment, all participants had repeated psychodiagnostic measurements to evaluate their symptoms at T_0, using self-reports (OQ45, BSI, BDI, ASI and BAI) and SCID-I interviews to assess diagnoses. Of these 150 patients 28 patients filled out T_1, with on average eight months (*M* = 242.1, *SD* = 105.1) between T_0 and T_1; 89 patients filled out T_2, with on average eleven months (*M* = 349.8, *SD* = 195.6) between T_0 and T_2; and 50 patients filled out T_3, with on average one and a half year (*M* = 553.6, *SD* = 187.4) between T_0 and T_3.

The prospective study

This study was executed between October 2013 and February 2014. Seventeen health care professionals (mostly psychologists) were interviewed by the master students who filled out the CSP (on paper) concerning their anxiety patients. As part of the standard procedure in their treatment, all participants had had psychodiagnostic measurements to evaluate their symptoms at T_0, using self-reports (OQ45, BSI, BDI, ASI and BAI) and SCID-I interviews to assess diagnoses. Patients had had between 0 and 4 treatment sessions at the moment of filling in the CPS. The master students made sure that all patients (that didn't drop out) filled in those same psychodiagnostic measurements at T_1, with on average three months (*M* = 88.7, *SD* = 45.8) between the first treatment session and the T_1 and four months (*M* = 121.7, *SD* = 45.4) between the T_0 and the T_1. At T_1 patients had between one and seven treatment sessions (*M* = 4 sessions).

Statistical analyses

The retrospective study

IBM SPSS Statistics 21 was used for all analyses. From the one hundred and fifty patients, there were 21 patients for whom we were unable to retrospectively fill out some CSP items on basis of their patient records because of missing data. Therefore CSP scores were calculated by using multiple data imputation. There were 37 variables used in this procedure, with 5 imputations, 100 parameter draws and 500 case draws: 4 CSP variables with missing data (effect of previous treatment, GAF score at premorbid functioning or intake, duration of symptoms or complaints, and course of the main disorder), 9 remaining CSP variables and 24

baseline or background variables that correlated significantly with those 4 CSP variables. A correlation matrix indicated that those 33 variables correlated significantly with those 4 missing CSP items, and could therefore be used to make the most accurate estimation of the missing data. This created a dataset of $N = 129$, in which all participants with missing values were removed, and a dataset of $N = 150$, in which all the missing values were substituted with imputed data. As shown in Appendix D, Tables 20 till 22, the data were randomly missing, and we therefore decided to only use the imputed data set in subsequent analyses.

To investigate the internal consistency of the CSP, Cronbach's alpha's were calculated on the items of the CSP. Further, to explore whether the items of the CSP were able to assign the anxiety disordered patients to the four categories of the CSP, one-way ANOVA's and a validity correlation matrix were used. For the ANOVA's each CSP category served as the independent variable, and the thirteen CSP items were the dependent variables. The validity of the CSP was calculated by means of a pooled correlation matrix in which the total scale of the adjusted CSP[2] was compared with the total scales and subscales of all the questionnaires included in the test battery at T_0.

Because of otherwise small sample sizes (and violations of assumptions), Kruskal-Wallis tests were performed, with the categories of the CSP as independent variable and the T_0-T_3 difference scores of the dependent variables as dependent variables, to analyze whether the category of the CSP predicts treatment progress. To investigate whether drop-out was predicted by the total score or the categories of the CSP, a binary logistic regression analysis was performed with the enter method and the non-drop-outs as reference category, with drop-out (yes/no) as dependent variable and separately the total score or the categories of the CSP as independent variable. The number of treatment drop-outs are calculated and the reasons why patients ended their treatment were investigated.

Prior to interpreting the results of the binary logistic regression analysis, several assumptions were evaluated in the retrospective study: both dependent and independent variables were normally distributed; the assumptions of normality, linearity and homoscedasticity of residuals were met; there were no univariate outliers or outliers in the XY-space; and multicollinearity would not interfere with our ability to interpret the outcome of the logistic regression analysis. Prior to interpreting the results of the ANOVA's, it was concluded that the assumptions of normality and normality of difference scores were not

[2] See 'Psychometric aspects of the CSP' in the Results.

violated. The assumption of homogeneity of variance was violated, but this is logical as the categories of the CSP should differ significantly.

The prospective study

IBM SPSS Statistics 21 was used for all analyses. Because of otherwise small sample sizes (and violations of assumptions), Kruskal-Wallis tests were performed, with the categories of the CSP as independent variable and the T_0-T_3 difference scores of the dependent variables as dependent variables, to analyze whether the category of the CSP predicts treatment progress. Because of the small sample size of the prospective study (N = 21), the small amount of drop-out (N = 2) and the selective drop-out (only patients from the second category of the CSP), a binary logistic regression analysis couldn't be performed to investigate whether drop-out was predicted by the total score or the categories of the CSP.

Prior to interpreting the results of the binary logistic regression analysis, several assumptions were evaluated in the retrospective study: both dependent and independent variables were normally distributed; the assumptions of normality, linearity and homoscedasticity of residuals were met; there were no univariate outliers or outliers in the XY-space; and multicollinearity would not interfere with our ability to interpret the outcome of the logistic regression analysis. Prior to interpreting the results of the ANOVA's, it was concluded that the assumptions of normality and normality of difference scores were not violated. The assumption of homogeneity of variance was violated, but this is logical as the categories of the CSP should differ significantly.

Results

Psychometric aspects of the CSP

With respect to the CSP we looked at Cronbach's alpha using the data of the 5th imputation, because Cronbach's alpha's of all imputations were almost identical. The checklist as a whole had an alpha of .566, which indicated a low internal consistency that was insufficient for research on group level (Evers, 2001). The column 'Cronbach's Alfa if Item Deleted' showed that the value of alpha could be increased by removing the items 'age' and 'level of education' respectively. After those removals, the CSP contained eleven items with an internal consistency that is sufficient for research on group level (α = .638, M = 13.56, SD = 4.67; Evers, 2001). See Table 11 for more details. The interpretation of the scores was adapted accordingly: total scores of the CSP subsequently ranged from 0 to 28, the first category ranged from 0 to 11 points, the second category from 12 to 17 points, the third category from 18 to 22 points and the fourth category from 23 to 28 points. See Appendix B1 (Dutch) and B2 (English) for the adjusted version of the CSP, that was used for all subsequent analyses.

Table 11

The Internal Consistency of CSP (in Cronbach's Alpha)

Removed item	Cronbach's Alpha[a]
Alpha with every item included	.566
Age	.629
Level of education	.638

[a]Cronbach's alpha for the 5th imputation.

One-way ANOVA's were used to compare the scores on individual items of the CSP between patients in the four categories of the CSP. Since the number of patients in the fourth category was too low (N = 4) to have sufficient statistical power, we decided to combine this category with the third category in subsequent analyses. The five items that differentiated significantly between all categories of the CSP were: comorbid depression or dysthymia, $F(2, 147) = 16.49$, $p < .001$, $\eta^2 = 0.18$, number of life areas affected by the anxiety disorder, $F(2, 147) = 43.43$, $p < .001$, $\eta^2 = 0.37$, duration of symptoms or complaints, $F(2, 147) = 28.85$, $p < .001$, $\eta^2 = 0.22$, number of previous treatments, $F(2, 147) = 18.95$, $p < .001$, $\eta^2 = 0.21$, and GAF score, $F(2, 147) = 27.62$, $p < .001$, $\eta^2 = 0.27$. The five items that differentiated

significantly between the first and second and first and third/fourth category of the CSP (subtle differentiation on low end) were: presence of multiple anxiety disorders, $F(2, 147) = 6.82$, $p < .01$, $\eta^2 = 0.08$, severity of the main anxiety disorder, $F(2, 147) = 13.35$, $p < .001$, $\eta^2 = 0.15$, number of situations the patient avoids, $F(2, 147) = 21.70$, $p < .001$, $\eta^2 = 0.23$, course of the main anxiety disorder, $F(2, 147) = 17.96$, $p < .001$, $\eta^2 = 0.20$, and effect of previous treatments, $F(2, 147) = 15.25$, $p < .001$, $\eta^2 = 0.17$. The item that differentiated significantly between the first and third/fourth and second and third/fourth category of the CSP (subtle differentiation on high end) was: number of non-anxiety comorbid disorders, $F(2, 147) = 9.77$, $p < .001$, $\eta^2 = 0.12$. Most of these differentiations had a small effect size, and the presence of multiple anxiety disorder had even an extremely small effect size. Only the number of life areas affected by the anxiety disorder and GAF score had a medium effect size.

The validity of the adjusted CSP was calculated by means of a pooled correlation matrix in which the total scale of the CSP was compared with the total scales and subscales of all the questionnaires included in the test battery at T_0. The CSP appeared to correlate significantly positive with the OQ45 total scale ($r = .403$, $p < .001$), BSI total scale ($r = .454$, $p < .001$), BSI depressive mood ($r = .419$, $p < .001$), BSI anxiety ($r = .371$, $p < .001$), BSI phobic anxiety ($r = .342$, $p < .001$), BDI total scale ($r = .447$, $p < .001$), ASI total scale ($r = .465$, $p < .001$). These correlations means that the CSP does not seems to generally represent symptom severity (including both anxiety and depression symptom severity) and is therefore not specific for the staging of anxiety disorders.

Descriptive statistics

The mean scores and standard deviations of the (sub)scales of all questionnaires are shown in Tables 7 (for the retrospective study) and 8 (for the prospective study).

Table 7

Means and Standard Deviations of the (Sub)scales of the Questionnaires (Retrospective Study; after data imputation)

	T_0 (*N* = 150)		T_1 (*N* = 28)		T_2 (*N* = 87)		T_3 (*N* = 150)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
CSP	16.43	4.79	--	--	--	--	--	--
OQ45 total scale	81.31	21.98	76.00	23.51	65.89	24.07	58.65	22.05
OQ45 symptom distress	47.79	12.64	43.82	15.23	37.57	14.86	32.37	12.42
BSI total scale	1.38	.67	1.20	.74	.94	.66	.65	.47
BSI anxiety	1.82	.89	1.41	1.03	1.12	.87	.82	.63
BSI phobic anxiety	1.53	.95	1.29	.85	.92	.75	.60	.62
BDI total scale	22.20	12.08	18.39	12.41	14.59	12.09	9.88	8.38

Table 8

Means and Standard Deviations of the (Sub)scales of the Questionnaires (Prospective Study)

	T_0 (*N* = 21)		T_1 (*N* = 19)	
	M	*SD*	*M*	*SD*
CSP	18.95	3.64	--	--
OQ45 total scale	80.57	15.53	77.79	13.17
OQ45 symptom distress	42.71	10.66	41.53	9.24
BSI total scale	59.29	30.55	51.74	27.44
BSI anxiety	1.51	.77	1.26	.52
BSI phobic anxiety	.87	.67	.62	.55
BDI total scale	21.38	11.66	18.95	8.95
BAI total scale[a]	36.55	8.77	33.29	8.10

[a]*N = 11,* [b]*N = 12,* [c]*N = 18*

General results with respect to treatment effect

As you can see in Figures 2 till 4 the overall scores on all the (sub)scales, anxiety measures as well as non-anxiety measures, improved in the course of treatment for patients in all CSP categories.

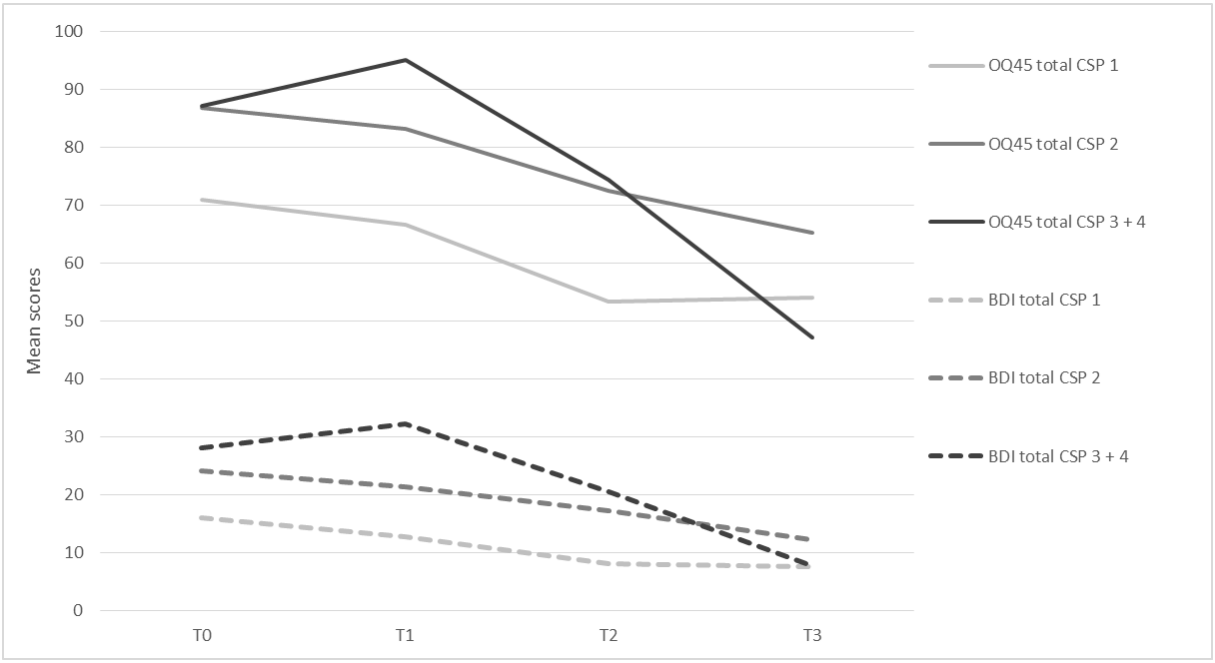

Figure 2. Treatment effect on the OQ45 and BDI for patients in all categories of the CSP.

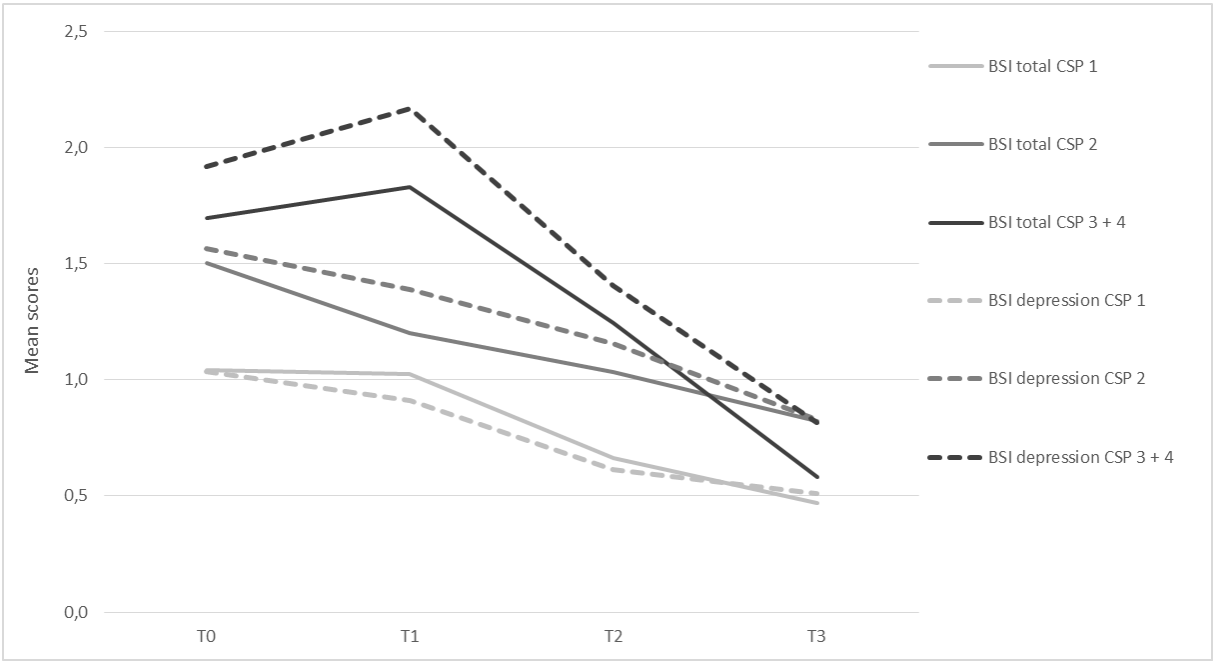

Figure 3. Treatment effect on the BSI total and depression score for patients in all categories of the CSP.

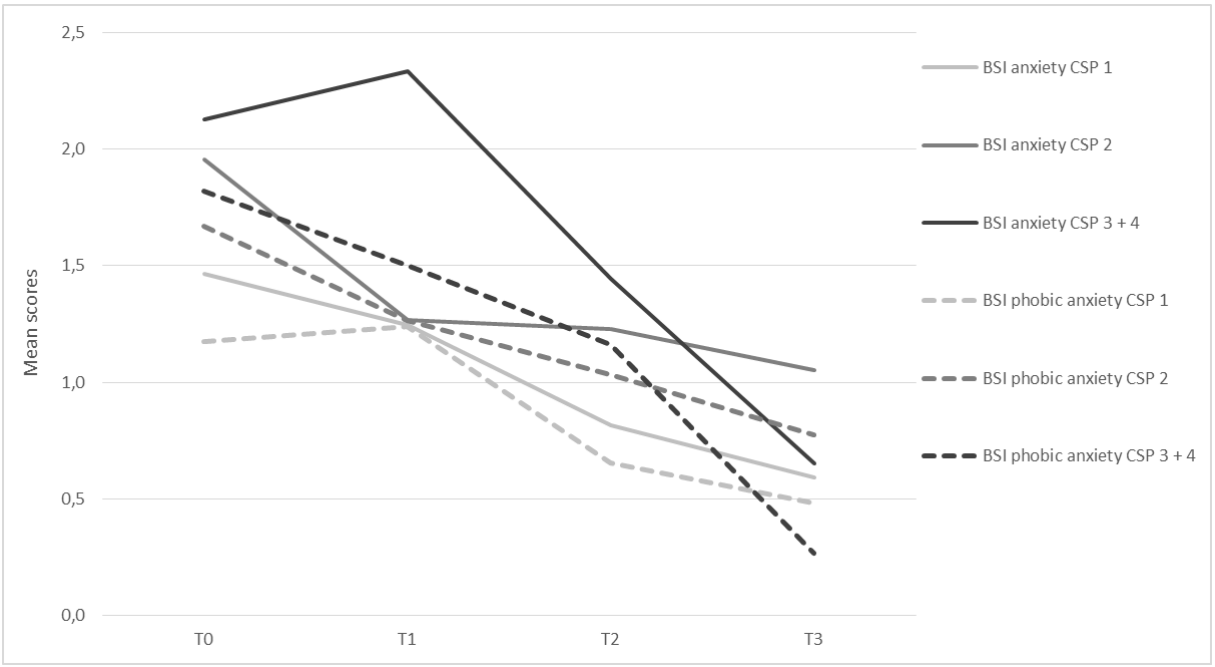

Figure 4. Treatment effect on the BSI anxiety scores for patients in all categories of the CSP.

A one-way ANOVA comparing symptom scores of the patients in the three CSP categories at baseline (table 7) showed that patients in the different CSP categories had significant different scores at baseline (T_0) on the total score of the OQ45, $F(2, 147) = 9.74$, $p < .001$, $\eta^2 = 0.12$; the total score of the BDI, $F(2, 147) = 12.76$, $p < .001$, $\eta^2 = 0.15$; the total score of the BSI, $F(2, 147) = 12.57$, $p < .001$, $\eta^2 = 0.15$; the depressive subscale of the BSI, $F(2, 147) = 12.41$, $p < .001$, $\eta^2 = 0.14$; the anxiety subscale of the BSI, $F(2, 147) = 7.69$, $p < .01$, $\eta^2 = 0.09$; and the phobic anxiety subscale of the BSI, $F(2, 147) = 5.79$, $p < .01$, $\eta^2 = 0.07$. In general, the scores of patients in the first CSP category at all times were lower than those of patients in the second category, and the scores of patients in the second category were lower than those of patients in the third/fourth category. It also seems that reduction of symptom severity across the categories was most prominent right after treatment (T_2) and at about six months after treatment (T_3).

After inspection of the graphs, it seems that, unlike patients in the first and second category of the CSP, patients in the third and fourth category regularly had little peaks at T_1, indicating some deterioration of scores before they improve their scores. Remarkably enough, when looking at the graphs, patients in the third category of the CSP subsequently seemed to experience the largest symptom reduction. A paired samples t-test between time points T_0 and T_3 indicated that patients in all categories of the CSP made significant treatment progress on all scores, but effect size increased as the category of the CSP became higher. That is, patients

in the second category had bigger effect sizes than patients in the first category and patients in the third/fourth category had bigger effect sizes than patients in the second category, which means that the more severely affected patients seemed to make relatively more treatment progress compared to the less severely affected patients. Note: all effect sizes showed at least a large effect.

Inductive statistics of the retrospective study (*N* = 150)

Hypothesis I

First, it was expected that the CSP predicts the course of treatment in patients with anxiety disorders. The categories of the CSP could significantly predict treatment progress between T_0 and T_3 measured by the difference scores on the total scale of the OQ45, $H(2) = 35.83$, $p < .001$; the total scale of the BSI, $H(2) = 35.99$, $p < .001$; the depressive subscale of the BSI, $H(2) = 33.79$, $p < .001$; the anxiety subscale of the BSI, $H(2) = 19.84$, $p < .001$; the phobic anxiety subscale of the BSI, $H(2) = 25.17$, $p < .001$; and the total scale of the BDI, $H(2) = 47.03$, $p < .001$. Jonckheere's test revealed a significant linear trend in the data: the higher the CSP category, the lower the (negative) median of the difference scores on the total scale of the OQ45, $J = 8710$, $z = -4.76$, $r = -.28$; the total scale of the BSI, $J = 8294$, $z = -5.34$, $r = -.32$; the depressive subscale of the BSI, $J = 8294$, $z = -5.57$, $r = -.33$; the anxiety subscale of the BSI, $J = 9761$, $z = -3.33$, $r = -.20$; the phobic anxiety subscale of the BSI, $J = 9068$, $z = -4.28$, $r = -.25$; and the total scale of the BDI, $J = 8304$, $z = -5.33$, $r = -.31$.These results indicate that patients in the higher categories of the CSP made significantly more treatment progress between T_0 and T_3 than patients in the lower categories of the CSP.

Hypothesis II

Second, it was expected that the CSP predicts drop-out in patients with anxiety disorders. Twenty-six patients dropped out of the retrospective study in total: 16 patients stopped showing up at treatment sessions and 10 patients stopped their treatment in agreement with their mental health care professionals. The percentage of drop-out was 11.3% for patients in the first category, 15.4% for the second category, 31.3% for the third/fourth category of the CSP[3].

[3] See Appendix E for all drop-out, including research drop-out.

As seen in Table 13, both the total score and the categories of the CSP could predict drop-out. The final logistic regression model (including the CSP) explained a significant amount of the original variability and therefore was a better fit than the original model (without the CSP). However, the pseudo R-squareds indicate that the model does only explain a very small percentage of the observed variation, with a very low odds ratio for the total scale of the CSP and a higher one for the categories of the CSP.

Table 13

Drop-out predicted by the CSP (Retrospective Study)

	B (*SE*)	Odds Ratio	95% BI
Total scale of the CSP [a]			
CSP total score	0.11 (0.05)*	1.11	[1.01, 1.23]
Constant	-3.07 (0.78)***	0.05	[0.01, 0.22]
Categories of the CSP [b]			
CSP category	0.63 (0.30)*	1.87	[1.03, 3.39]
Constant	-2.79 (0.67)***	0.06	[0.02, 0.23]

Note. The pooled results were used for all data in this table.

[a] For the 5th imputation: R^2 = .03 (Cox & Snell), .05 (Nagelkerke). Model $\chi^2(1)= 4.97$, $p < .05$.

[b] For the 5th imputation: R^2 = .03 (Cox & Snell), .05 (Nagelkerke). Model $\chi^2(1)= 4.95$, $p < .05$.

*$p < .05$. **$p < .01$. *** $p < .001$.

Furthermore, a frequency analysis of drop-outs showed that patients with low GAF scores at baseline between 40 and 49 were three times more likely to drop-out and that drop-outs, in comparison with the entire group of patients that participated in this study, were two times more likely to have PTSD, were a little bit older (30 till 49 year instead of 20 till 39 year), were one and a half times more likely to be a (first generation) migrant, were two and a half times more likely to be of Moroccan descent, were one and a half times more likely to neither work nor study and were one and a half times more likely to have children and live with children (as a single parent or with a partner).

Inductive statistics of the prospective study

Hypothesis I

First, it was expected that the CSP predicts the course of treatment in patients with anxiety disorders. The categories of the CSP couldn't significantly predict treatment progress between

T_0 and T_1 measured by the difference scores on the total scale of the OQ45, $H(2) = 1.48$, $p > .05$; the total scale of the BSI, $H(2) = 1.83$, $p > .05$; the depressive subscale of the BSI, $H(2) = 0.81$, $p > .05$; the anxiety subscale of the BSI, $H(2) = 0.17$, $p > .05$; the phobic anxiety subscale of the BSI, $H(2) = 0.34$, $p > .05$; or the total scale of the BDI, $H(2) = 1.71$, $p > .05$.

Hypothesis 2: drop-out could not be predicted due to the low sample size of the study.

Discussion

In this study we aimed to investigate the relative contribution of the items of the CSP to the correct prediction of the illness stage of a patient's anxiety disorder. Next, it was expected that the CSP predicts the course of treatment in patients with anxiety disorders, and it also predicts drop-out in patients with anxiety disorders, in the following way: patients with high scores on the CSP (e.g. the third and fourth category) would be less likely to make progress in their treatment and be more likely to drop out, because the first category would contain the patients with the least illness severity and the fourth category would contain patients with the highest illness severity.

The original CSP (Appendix A) had a low internal consistency and was therefore insufficient for research on group level (Evers, 2001). After removal of the items 'age' and 'level of education' the CSP (Appendix B) had an internal consistency which was sufficient for research on group level (Evers, 2001). This version was used for further analyses. Most items of the CSP contributed to a significant differentiation between the categories of the CSP: five items differentiated between all categories, five items between the first and second and first and third/fourth category, and one item between the first and third/fourth and second and third/fourth category. However, the CSP did not seem to be specific for staging in anxiety disorders, since it also correlated significantly positive with non-anxiety measures. This seems logical, considering the high co-occurrence between anxiety disorders and depression, and considering the fact that staging with the CSP is closely associated with symptom severity.

These first results established that patients in the different CSP categories had significantly different baseline (T_0) scores on all (sub)scales. Second, at all times (except at T_3), the scores of patients in the first CSP category were significantly lower than those of patients in the other categories, and there was a significant reduction of symptom severity for all CSP categories right after treatment (T_2), indicating that the more severely affected patients made stronger improvements after treatment ended (at T_3), than the less severely affected patients (who continued to improve a little).

Third, the results showed that patients in all categories of the CSP made significant treatment progress (between T_0 and T_3) on all scores, anxiety measures as well as non-anxiety measures; and, interestingly effect size increased as the category of the CSP became higher, which means that the more severely affected patients made relatively more treatment progress

compared to less severely affected patients. In the retrospective study, the categories of the CSP could significantly predict treatment progress (between T_0 and T_3) measured by the difference scores on all (sub)scales. There was a linear trend in the data: the higher the CSP category, the greater the treatment progress. Patients in the third and fourth category of the CSP seemed to experience the largest symptom reduction. This was in contrast to what was expected by our hypotheses, i.e. that the patients in the more severely affected CSP categories (3 and 4) would benefit less from treatment than those in the less severely affected ones. Thus, the CSP did predict the course of treatment in patients with anxiety disorders but in a different direction than hypothesized. Nevertheless, these severely affected patients seemed to experience the largest symptom reduction, which means that not all patients with severe pretreatment psychopathology or comorbidity respond poorly to treatment. In fact, the magnitude of reduction of scores can be identical between people with severe versus milder anxiety disorders, although people with initially more severe symptoms could be more likely to have clinically significant residual symptoms at the end of treatment (Taylor, Abramowitz & McKay, 2012). In the prospective study, however, the categories of the CSP could not significantly predict treatment progress between T_0 and T_1 measured by any of the difference scores. This could be because of low statistical power, due to the small sample size, and because of the small timespan of the prospective study: on average, there were only three months between the first treatment session and T_1 (instead of the one and a half year in the retrospective study), in which patients only had had up to seven treatment sessions (with an average of four).

Further, unlike patients in the first and second category of the CSP, patients in the third/fourth category regularly had some deterioration of scores at T_1 before they improved their scores at T_2 and T_3. One possibility to explain this interesting finding could be that this more severely affected group has more difficulty in following treatment at first, due to ingrained behavioral patterns (because of long duration and chronic course of the disorder), negative treatment expectation as a result of previous failed treatments, or more often a comorbid depressive disorder that first needed to be addressed.

Fourth, it was expected that the CSP predicts drop-out in patients with anxiety disorders. The results showed that both the total score and the categories of the CSP could predict drop-out, but only a very small percentage of the observed variation could be explained by the model. When examining the drop-out rates per category, most patients who dropped out belonged to the third/fourth category (31.3%), followed by patients in the second

category (15.4%) and the lowest proportion of dropouts in the first category (11.3%) of the CSP. This pattern was exactly like it was predicted to be: the higher the category, the more likely to drop-out. This means that special attention should be given to the more severely affected patients in the third CSP category to prevent premature drop-out. In our opinion, instead of dropping out, these patients often benefit more from an extended course of treatment or a more intensive treatment, although at the same time the greater demands of intensive treatment might increase the risk of treatment drop-out (Taylor, Abramowitz & McKay, 2012). It appeared that patients with a lower GAF score at baseline (indicating low levels of overall functioning) were three times more likely to drop-out and that drop-outs, in comparison with the entire group of patients that participated in this study, were two times more likely to have PTSD, were somewhat older, were more likely to be a (first generation) migrant, were more likely to be of Moroccan descent, were more likely to neither work nor study and were more likely to have children to look after. The fact that patients diagnosed with PTSD dropped out more often (more than a quarter) is in line with the literature, as traumatization is a known risk factor for completing treatment and treatment typically involves exposure to trauma-related stimuli and imagery, which can be aversive: the majority of patients receiving psychotherapy tend to drop out before they have received an "adequate dose" for symptom relief (Barrett, Chua, Crits-Cristoph, Gibbons, & Thompson, 2008; Cully, et al., 2008; as cited in Angeli, 2009). The older age of patients who dropped out might be explained by the fact that older patients might have undergone more often previous treatments without success or with rapid recurrence of symptoms. As a result they might be demoralized and prematurely drop-out. Another reason could be that older patients maintain their treatment gains less well than younger patients (Foa, et al., 1983) and are less flexible to change. This could be a burden for therapy, as they suddenly have to learn how to change their ways and change isn't that easy to accomplish anymore. The fact that patients with a low level of functioning at the start of treatment were more likely to drop-out means that, from the start, less compensation strategies are available, and low GAF scores may indicate worse premorbid functioning as well, which may mean that patients have less room to improve functioning in general with symptom relief. As a consequence, significant attention should be paid to this aspect during treatment, apart from obtaining symptom reduction. With respect to migrant status (Maroccan descent), Sue, McKinney and Allen (1976) found that ethnicity was a very important predictor in early termination of treatment and in length of therapy. Although they primarily investigated this in Afro-American patients, this could be true for other minorities as well. Ouellet-Plamondon, Rousseau, Nicole and Abdel-Baki (2015), and

also a study in our own AAA sample (Rijkeboer, et al., in review), found that (first- and second-generation) immigrants were more likely than non-immigrants to disengage from treatment. In an overview of the current literature Jonsdottir and Waghorn (2015) found that the proportions of people employed decreased with the more severe disorder categories, indicating that severe psychiatric illnesses might contribute to employment struggles for people with these illnesses, across countries. However, the causality might also point into the opposite direction, i.e. that work struggles cause psychological problems. Issakidis and Andrews (2004) found that the fact of having at least one child also heightens the probability of drop-out, which could explain why drop-outs, more often than completers, had children.

There could also be other explanations for drop-out that have not been measured in our research. Hofmann and Suvak (2006), for example, found that drop-outs rated the treatment rationale as less logical than completers at the beginning of treatment. Non-adherence with treatment has also been found to be related to patients' motivation for treatment (Kortrijk, et al, 2012) and their expectations and opinions about treatment (Santana & Fontenelle, 2011). Therefore, it could be interesting to pay more attention to these variables – as part of the CSP – before starting treatment. Social support from their loved ones and/or from their employer, or educational institution, is also important for patients to complete their treatment with confidence. Maybe all these variables could be added to the CSP at a later stage to enhance its predictive value.

Ultimately, we can conclude that patients generally did improve significantly, and that the CSP was able to predict this. Besides the recommendations mentioned here-above about adding extra variables to predict drop-out, we like to make some extra suggestions for further improvement of the CSP. First, the fact that age wasn't a good predictor, could be explained by the fact that this was the age at intake and not the age at onset of the anxiety disorder. It was thought that an early age of onset (young patient) would predict a bad prognosis for their anxiety disorder. The older patients often have a longer duration of their anxiety disorder (more than ten years), which often means their age of onset was early on in life. Illness duration was included in the CSP. However, this item does not fully cover whether a patient has a young age at onset, and therefore it might add to predictability of the CSP when age at onset would be specifically added to the CSP. Interestingly, the level of education could only marginally differentiate between the first and third category of the CSP. This is in line with the literature, because level of education often influences other variables indirectly, which in turn affect the severity of the anxiety disorder. Alonso, et al. (2004), for example, found that a

higher educational level was associated with a higher risk of pure anxiety disorder, whereas a low educational level was associated with a higher risk for anxiety-comorbid mood disorder. This is why we already deleted this item from the CSP.

Second, there are some psychometric properties of the CSP that can be improved. The fact that many items couldn't significantly differentiate between the second and third category of the CSP could be due to the scoring of these items. The number of affected life areas, duration of anxiety symptoms and course of the anxiety disorder all have a limited score range (0-2). The patients of the AAA might all have scores on these items that easily reach this maximum score, adding to a ceiling effect of the score. Maybe if there were more options to choose from, the more severely affected patients (third and fourth category) would score higher than patients who scored in the second category, and these items would become more sensitive to pick up differences between CSP categories. The fact that the number of non-anxiety comorbid disorders couldn't significantly differentiate between the first and second category of the CSP could simply be due to the fact that comorbidity occurs by far more often in the more severely affected patients, like the ones in the third and fourth category of the CSP. In the first category of the CSP around 83% of the patients had no comorbid disorder, as had 76% of the patients in the second category, whereas patients in the third and fourth category had no comorbid disorder in only 42% of cases. So the number of non-anxiety comorbid disorders can indicate the most severe patients, but cannot differentiate between the first and second category of the CSP.

Therefore it's important to remark that scoring possibilities could be adjusted to increase the internal consistency of the CSP even more, increase the possibility that patients score in the fourth category of the CSP, and that patients are distributed more accurately between all categories. In this study most patients scored in the first, second and third category, and only four patients scored in the fourth category, although we would expect this patient population of Altrecht to score especially in the second, third and fourth category, as this is a specialized mental health institution for the more severe cases of anxiety disorders. Therefore I would like to suggest to change the future interpretation of the scores so it matches the population of the AAA. See Appendix F1 (Dutch) and F2 (English) for these suggested changes. A subsequent suggestion would be to then re-test the CSP on its staging properties using anxiety patients in the basic mental health care as well as in specialized mental health care, to see if the CSP discriminates and if internal consistency and categorization of the CSP is improved.

Conclusively, the CSP seems a reliable checklist which differentiates accurately between the different stages of anxiety disorders. It can predict the course of treatment in patients with anxiety disorders and also which patients (e.g. the ones in the higher categories) will drop-out if no appropriate measures are taken (e.g. more intensive treatment). However, it needs some more development and fine-tuning.

Literature

Alonso, J., Angermeyer, M.C., Bernert, S., Bruffaerts, R., Brugha, T.S., Bryson, H., de Girolamo, G., de Graaf, R., Demyttenaere, K., Gasquet, I., Haro, J.M., Katz, S.J., Kessler, R.C., Kovess, V., Lépine, J.P., Ormel, J., Polidori, G., Russo, L.J., Vilagut, G., Almansa, J., Arbabzadeh-Bouchez, S., Autonell, J., Bernal, M., Buist-Bouwman, M.A., Codony, M., Domingo-Salvany, A., Ferrer, M., Joo, S.S., Martínez-Alonso, M., Matschinger, H., Mazzi, F., Morgan, Z., Morosini, P., Palacín, C., Romera, B., Taub, N. & Vollebergh, W.A.M. (2004). 12-Month comorbidity patterns and associated factors in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatrica Scandinavica, 109(420)*, 28-37.

Altamura, A.C., Buoli, M., Albano, A. & Dell'Osso, B. (2010). Age at onset and latency to treatment (duration of untreated illness) in patients with mood and anxiety disorders: a naturalistic study. *International Clinical Psychopharmacology, 25*, 172-179.

Altamura, A.C., Santini, A., Salvadori, D. & Mundo, E. (2005). Duration of untreated illness in panic disorder: a poor outcome risk factor? *Neuropsychiatric Disease and Treatment, 1(4)*, 345–347.

American Psychiatric Association (2001). *DSM-IV-TR.* Pearson: Amsterdam.

Angeli, N.C. (2009). *Adherence to psychotherapy for post-traumatic stress in veterans of military combat in Afghanistan (Operation Enduring Freedom) and Iraq (Operation Iraqi Freedom).* Atlanta: Georgia State University.

Beck, A.T., Epstein, N., Brown, G. & Steer, R.A. (1988). An inventory for measuring clinical anxiety: psychometric properties. *Journal of Consulting and Clinical Psychology, 56(6)*, 893-897.

Beck, A.T., Ward, C.H., Mendelson, M., Mock, J. & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561-571.

Bet, P.M., Hugtenburg, J.G., Penninx, B.W.J.H., Van Balkom, A., Nolen, W.A. & Hoogendijk, W.J.G. (2013). Treatment inadequacy in primary and specialized care patients with depressive and/or anxiety disorders. *Psychiatry Research, 210(2)*, 594-600.

Christiana, J.M., Gilman, S.E., Guardino, M., Mickelson, K., Morselli, P.L., Olfson, M. & Kessler, R.C. (2000). Duration between onset and time of obtaining initial treatment among people with anxiety and mood disorders: an international survey of members of mental health patient advocate groups. *Psychological Medicine, 30*, 693–703.

Cohen J. E. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc

De Beurs, E. & Zitman, F.G. (2006). De Brief Symptom Inventory (BSI): de betrouwbaarheid en validiteit van een handzaam alternatief voor de SCL-90. [The Brief Symptom Inventory (BSI): the reliability and validity of a brief alternative of the SCL-90.] *Maandblad Geestelijke Volksgezondheid, 61*, 120-141.

De Graaf, R., Ten Have, M. & Van Dorsselaer, S. (2010). *De Psychische gezondheid van de Nederlandse bevolking. NEMESIS-2: opzet en eerste resultaten.* Utrecht: Trimbos Instituut.

Derogatis, L.R. (1975). *Brief Symptom Inventory.* Baltimore, MD: Clinical Psychometric Research.

Ebert, D.D., Gollwitzer, M., Riper, H., Cuijpers, P., Baumeister, H. & Berking, M. (2013). For whom does it work? Moderators of outcome on the effect of a transdiagnostic internet-based maintenance treatment after inpatient psychotherapy: randomized controlled trial. *Journal of Medical Internet Research, 15(10)*, 231-247.

Evers, A. (2001). *Beoordelingssysteem voor de kwaliteit van tests.* Amsterdam: COTAN/ Nederlands Instituut van Psychologen.

Fernández, A, Haro, J.M., Martinez-Alonso, M, Demyttenaere, K., Brugha, T.S., Autonell, J., De Girolamo, G., Bernert, S., Lépine, J.P. & Alonso, J. (2007). Treatment adequacy for anxiety and depressive disorders in six European countries. *British Journal of Psychiatry, 190*, 172-173.

Foa, E.B., Grayson, J.B., Steketee, G.S., Doppelt, H.G., Turner, R.M. & Latimer, P.R. (1983). Succes and failure in the behavioral treatment of obsessive-compulsives. *Journal of Consulting & Clinical Psychology, 51(2)*, 287-297.

Høifødt, R.S., Strøm, C., Kolstrup, N., Eisemann, M. & Waterloo, K. (2011). Effectiveness of cognitive behavioural therapy in primary health care: a review. *Family Practice, 28*, 489-504.

Jonsdottir, A. & Waghorn, G. (2015). Psychiatric disorders and labour force activity. *Mental Health Review Journal, 20(1)*, 13-27.

Kessler, R.C., Ruscio, A.M., Shear, K. & Wittchen, H.-U. (2009). Behavioral neurobiology of anxiety and its treatment. In Stein, M.B. & Steckler, T. (eds.), *Current topics in behavioral neurosciences 2* (pp. 21-35). Berlin: Springer-Verlag Berlin Heidelberg.

Kortrijk, H.E., Mulder, C.L., Drukker, M., Wiersma, D. & Duivenvoorden, H.J. (2012). Duration of assertive community treatment and the interpretation of routine outcome data. *Australian and New Zealand Journal of Psychiatry, 46(3),* 240-248.

McLean, P.D., Whittal, M.L., Thordarson, D.S., Taylor, S., Soechtig, I., Koch, W.J., Paterson, R. & Anderson, K.W. (2001). Cognitive versus behaviour therapy in the group treatment of obsessive-compulsive disorder. *Journal of Consulting and Clinical Psychology, 69(2),* 205–214.

Mululo, S.C.C., Bezerra de Menezes, G., Vigne, P. & Fontenelle, L.F. (2012). A review on predictors of treatment outcome in social anxiety disorder. *Revista Brasileira de Psiquiatria, 34,* 92-100.

Nunnally, J.C. & Bernstein, I.H. (1994). *Psychometric theory.* New York: McGraw-Hill.

Ouellet-Plamondon, C., Rousseau, C., Nicole, L. & Abdel-Baki, A. (2015). Engaging immigrants in early psychosis treatment: a clinical challenge. *Psychiatric Services, 66(3),* n.p.

Reiss, S., Peterson, R.A., Gursky, D.M. & McNally, R.J. (1986). Anxiety sensitivity, anxiety frequency, and the prediction of fearfulness. *Behaviour Research and Therapy, 24,* 1-8.

Rijkeboer, M.M. et al. (in review).

Santana, L. & Fontenelle, L.F. (2011). A review of studies concerning treatment adherence of patients with anxiety disorders. *Patient Preference and Adherence, 5,* 427–439.

Slaap, B.R., Van Vliet, I.M., Westenberg, H.G.M. & Den Boer, J.A. (1996). MHPG and heart rate as correlates of nonresponse to drug therapy in panic disorder patients. *Psychopharmacol, 127,* 353-358.

Smith, J.D., Van Ryzin, M.J., Fowler, J.C. & Handler, L. (2014). Predicting response to intensive multimodal inpatient treatment: a comparison of single- and multiple-class growth modeling approaches. *Journal of Personality Assessment, 96(3),* 306-315.

Somers, J.M., Goldner, E.M., Waraich, P. & Hsu, L. (2006). Prevalence and incidence studies of anxiety disorders: a systematic review of the literature. *Canadian Journal of Psychiatry, 51,* 100-113.

Steer, R.A., Beck, A.T. & Garrison, B. (1986). Applications of the Beck Depression Inventory. In N. Sartorious & T.A. Ban (Eds.), *Assessment of depression* (pp. 121-142). Geneva: World Health Organization.

Sue, S., McKinney, H., & Allen, D.B. (1976). Predictors of the duration of therapy for clients in the community mental health system. *Community Mental Health Journal*, *12*, 365-375.

Taylor, S., Abramowitz, J.S. & McKay, D. (2012). Non-adherence and non-response in the treatment of anxiety disorders. *Journal of Anxiety Disorders, 26*, 583–589.

Taylor, S., Zvolensky, M., Cox, B., Deacon, B., Heimberg, R., Ledley, D.R., Abramowitz, J.S., Holaway, R.M., Sandin, B., Stewart, S.H., Coles, M., Eng, W., Daly, E.S., Arrindell, W.A., Bouvard, M. & Cardenas, S.J. (2007). Robust dimensions of anxiety sensitivity: development and initial validation of the Anxiety Sensitivity Index-3 (ASI-3). *Psychological Assessment, 19*, 176-188.

Tolin, D.F., Gilliam, C.M. & Dufresne, D. (2010). The economic and social burden of anxiety disorders. In Stein, D.J., Hollander, E. & Rothbaum, B.O. (Eds.), *Textbook of anxiety disorders* (2nd ed., pp. 731–746). Arlington, VA: American Psychiatric Publishing.

Umphress, V.J., Lambert, M.J., Smart, D.W., Barlow, S.H., Clouse, G. (1997). Concurrent and construct validity of the outcome questionnaire. *Journal of Psychoeducational Assessment, 15(1)*, 40-55.

Van Balkom, A.L.J.M., Oosterbaan, D.B., Batelaan, N., Cath, D.C., Hendriks, G.-J., Van Megen, H.J.G.M., Schruers, K., Spijker, J., Van der Wee, N. & Van Vliet, I.M. (2012). De karakterisering van angststoornissen: stageren en profileren met gezond verstand. *Tijdschrift Voor Psychiatrie, 54*, 935-940.

Van Balkom, A.L.J.M., Van Vliet, I.M., Emmelkamp, P.M.G., Bockting, C.L.H., Spijker, J., Hermens, M.L.M., Meeuwissen, J.A.C., namens de Werkgroep Multidisciplinaire richtlijnontwikkeling Angststoornissen/Depressie [on behalf of the Workgroup Multidisciplinary Guideline Development Anxiety Disorders/Depression] (2013). Multidisciplinaire richtlijn Angststoornissen (Derde revisie). Richtlijn voor de diagnostiek, behandeling en begeleiding van volwassen patiënten met een angststoornis. [Multidisciplinary guideline Anxiety Disorders (Third revision). Guideline for the diagnostics, treatment and guidance of adult patients with an anxiety disorder.] Utrecht: Trimbos-instituut.

Wang, P.S., Aguilar-Gaxiola, S., AlHamzawi, A.O., Alonso, J., Andrade, L.H., Angermeyer, M., Borges, G., Bromet E.J., Bruffaerts, R., Bunting, B., Caldas de Almeida, J.M., Florescu, S., De Girolamo, G., De Graaf, R., Gureje, O., Haro, J.M., Hinkov, H.R.,

Hum C.Y., Karam, E.G., Kovess, V., Lee, S., Levinson, D., Ono, Y., Petukhova, M., Posada-Villa, J., Sagar, R., Seedat, S., Wells, J.E., Kessler, R.C. (2011). Treated and untreated prevalence of mental disorder worldwide. In Thornicroft, G., Szmukler, G. & Mueser, K.T. (Eds.), *Oxford textbook of community mental health* (pp. 50-66). New York: Oxford University Press.

Wells, M.G, Burlingame, G.M, Lambert, M.J., Hoag, M.J. & Hope, C.A. (1996). Conceptualization and measurement of patient change during psychotherapy: development of the outcome questionnaire and youth outcome questionnaire. *Psychotherapy, 33(2),* 275-283.

White, K.S., Allen, L.B., Barlow, D.H., Gorman, J.M., Shear, M.K. & Woods, S.W. (2010). Attrition in a multicenter clinical trial for panic disorder. *Journal of Nervous and Mental Disease*, 198, 665-671.