



UvA-DARE (Digital Academic Repository)

Concordance between self-reported and observer-rated anxiety severity in outpatients with anxiety disorders

The Leiden routine outcome monitoring study

Schat, A.; van Noorden, M.S.; Giltay, E.J.; Noom, M.J.; Vermeiren, R.R.J.M.; Zitman, F.G.

DOI

[10.1111/papt.12134](https://doi.org/10.1111/papt.12134)

Publication date

2017

Document Version

Final published version

Published in

Psychology and Psychotherapy

[Link to publication](#)

Citation for published version (APA):

Schat, A., van Noorden, M. S., Giltay, E. J., Noom, M. J., Vermeiren, R. R. J. M., & Zitman, F. G. (2017). Concordance between self-reported and observer-rated anxiety severity in outpatients with anxiety disorders: The Leiden routine outcome monitoring study. *Psychology and Psychotherapy*, *90*(4), 705-719. <https://doi.org/10.1111/papt.12134>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)



Concordance between self-reported and observer-rated anxiety severity in outpatients with anxiety disorders: The Leiden routine outcome monitoring study

Anke Schat^{1*}, Martijn S. van Noorden¹, Erik J. Giltay¹,
Marc J. Noom^{2,3}, Robert R.J.M. Vermeiren⁴ and Frans G. Zitman¹

¹Department of Psychiatry, Leiden University Medical Center, The Netherlands

²Department of Psychiatry, Zaans Medisch Centrum, The Netherlands

³Department of Child Development and Education, University of Amsterdam, Amsterdam, The Netherlands

⁴Department of Child and Adolescent Psychiatry, Curium-Leiden University Medical Center, Oegstgeest, The Netherlands

Objectives. Anxiety severity measures can be self-report or observer-rated. Although mostly these measures concur, they can diverge markedly. We examined concordance between two anxiety scales: the observer-rated Brief Anxiety Scale (BAS) and the self-report Brief Symptom Inventory 12-item version (BSI-12), and described associations between patient characteristics and discordance.

Design. The study used an observational design, using prospective data from 2,007 outpatients with DSM-IV-TR panic disorder with or without agoraphobia, agoraphobia without panic, social phobia, and/or generalized anxiety disorder.

Methods. Overall agreement was described using Pearson's product-moment correlation coefficient. Associations between patient characteristics and discordance (defined as $|Z\text{-BAS} - Z\text{-BSI-12}| \geq 1$) were evaluated with univariable and multivariable multinomial logistic regression analyses.

Results. Overall correlation between BAS and BSI-12 was positive and strong ($r = .59$). Discordance occurred in 24.8% of patients ($[Z\text{-BAS} \geq Z\text{-BSI-12} + 1] = 12.2\%$; $[Z\text{-BAS} \leq Z\text{-BSI-12} - 1] = 12.6\%$). Patients with higher observed than self-reported anxiety severity did not differ from concordant patients. Patients with lower observed than self-reported anxiety severity more often had panic disorder, less often had social phobia, and had higher scores on cluster B and C personality characteristics than concordant patients. Lower observed than self-reported anxiety severity was best predicted by panic disorder, social phobia, and affective lability.

Conclusions. Results demonstrate that the use of a single source of information gives a one-sided view of pathology. A multimethod approach is highly preferable, as this allows for assessment across different domains and through multiple sources of information, and as such, provides clinicians with vital information.

*Correspondence should be addressed to Anke Schat, Department of Psychiatry, Leiden University Medical Center, PO Box 7500 2300 RC Leiden, the Netherlands (email: a.schat@lumc.nl).

Practitioner points

- When assessing anxiety severity, the use of self-report measures provides additional information to observer-rated measures.
- In patients who have strong cluster B and C personality traits, anxiety severity might be overlooked, even by trained observers.
- The use of a multimethod assessment strategy is preferable in anxiety severity assessment.

When quantifying the severity of psychiatric disorders, clinicians rely on psychiatric rating scales. These scales can be either observer-rated or self-report. Observer-rated instruments are often regarded as the primary source of information (Hamilton, 1976; Moller, 2000). Self-report measures, on the other hand, are more efficient in terms of time and costs and are therefore increasing in popularity in clinical practice, where the need to quantify symptom severity competes with the need to economize. Interestingly, several studies have demonstrated that observer-rated severity does not necessarily correspond to patient-reported severity. Previous research on depression rating scales demonstrated that correlations between observer and self-rated severity ranged from .12 (Dorz, Borgherini, Conforti, Scarso, & Magni, 2004), .28 (Dunlop *et al.*, 2011), .40 (Enns, Larsen, & Cox, 2000), .45 (Rane *et al.*, 2010), and .46 (Dunlop *et al.*, 2011), to .59 and .60 (Carter, Frampton, Mulder, Luty, & Joyce, 2010). Although repeated measures, even by the same rater, will always show some discrepancy due to random error, the discrepancy between observer and self-reported severity seems to consist of more than random error, as it has been linked to several patient characteristics, such as age (Carter *et al.*, 2010) (Dorz *et al.*, 2004; Enns *et al.*, 2000), gender (Carter *et al.*, 2010; Jolly, Wiesner, Wherry, Jolly, & Dykman, 1994), education level (Enns *et al.*, 2000), marital status (Dorz *et al.*, 2004), and psychiatric history (Dorz *et al.*, 2004). In addition, the level of concordance between observer-rated and self-report instruments appears to be related to personality. Lower scores on observer-rated relative to self-report instruments were associated with more personality disorder in depression (Dorz *et al.*, 2004; Rane *et al.*, 2010), as well as high neuroticism, low extraversion, low agreeableness (Enns *et al.*, 2000), high novelty seeking, and high reward dependence (Carter *et al.*, 2010).

The majority of these previous studies have focussed on depression (Carter *et al.*, 2010; Dorz *et al.*, 2004; Dunlop *et al.*, 2011; Enns *et al.*, 2000; Rane *et al.*, 2010), with only one study describing concordance in generalized anxiety disorder (Hopko *et al.*, 2000), one study focussing on patients with symptoms of anxiety and depression in primary care (Lubaczewski, Shepherd, Fayyad, & Guico-Pabia, 2014), and one study of concordance in adolescent inpatients with a variety of diagnoses (Jolly *et al.*, 1994). In addition, although associations between discordance and the presence of personality disorders as well as NEO five personality factors have been studied, no in-depth exploration of personality and concordance exists. We therefore conducted a study of concordance between observer-rated and self-report instruments measuring anxiety severity in a large naturalistic sample of outpatients with commonly occurring anxiety disorders (panic disorder with or without agoraphobia, agoraphobia without panic, social phobia, and generalized anxiety disorder) with special focus on associations with personality traits. We set out to describe concordance between an observer-rated measure of anxiety severity, the Brief Anxiety Scale (BAS; Tyrer, Owen, & Cicchetti, 1984), and a self-report measure of anxiety severity (the Brief Symptom Inventory 12-item version [BSI-12], De Beurs & Zitman, 2006; Roy-Byrne *et al.*, 2010), on group level. Next, we examined associations between concordance and a set of patient characteristics that were previously associated with concordance in depression, as well as an extensive set of

personality traits. Finally, we analysed which patient characteristics best predicted discordance. Results may help to weigh the benefits and costs of self-report and observer-rated measures in quantifying anxiety severity.

Methods

Participants

All subjects were outpatients at the department of psychiatry of a university medical centre or an affiliated regional mental health care provider. Within both centres, as part of clinical practice, all patients were administered an extensive battery of diagnostic and psychometric measures by trained research nurses and through supervised computerized self-report. This happened at baseline and approximately every 3 months of follow-up although for the purpose of this study we used only the baseline data. This procedure is known as routine outcome monitoring (ROM) and is described in more detail (De Beurs *et al.*, 2011). Patients were aged 18 through 65 and had adequate command of the Dutch language. All patients had been referred by their general practitioner for the treatment of mood, anxiety, or somatoform disorders and met DSM-IV-TR diagnostic criteria for at least one of the following disorders: panic disorder with or without agoraphobia, agoraphobia without panic, social phobia, and generalized anxiety disorder (allowing for comorbid mood or somatoform disorders). The routine outcome monitoring infrastructure holds computerized administration of questionnaires, and this does not allow respondents to skip single items. When an item is not answered, the patient cannot progress to the next item and the questionnaire cannot be completed. If a questionnaire is ended without completing it, answers cannot be saved. Patients who had missing data resulting from the incidental failure to administer complete questionnaires, or who had a large time gap (more than 21 days) between the administration of questionnaires, were excluded from the analyses.

Measures

Anxiety severity

Observer-rated anxiety severity was assessed using the BAS (Tyrer *et al.*, 1984). The BAS is a 10-item observer-rated scale derived from the abbreviated Comprehensive Psychopathological Rating Scale (CPRS, Asberg, Montgomery, Perris, Schalling, & Sedvall, 1978; Goekoop *et al.*, 1991). The total score equals the sum score of all 10 items on a 7-point Likert scale (0–6, range 0–60). It has adequate internal consistency with Cronbach's alpha of .64 in our cohort. It assesses the main components of all anxiety disorders, covering psychic and somatic components. Patient-reported severity of anxiety symptoms was measured at intake with the Dutch version of the BSI-12 (De Beurs & Zitman, 2006; Roy-Byrne *et al.*, 2010). The BSI-12 is a self-report measure comprising items of the anxiety and somatization subscales of the 18-item version of the BSI, which has in turn been derived from the BSI (Zabora *et al.*, 2001). The total score equals the sum score of 12 items on a 5-point Likert scale (0–4, range 0–48). It has good internal consistency with Cronbach's alphas between .79 and .84 (Franke *et al.*, 2011). Internal consistency (i.e., Cronbach's alpha) in our cohort was .90. On both scales, a higher score corresponds to more severe anxiety. Table 4 lists the items comprising both the BAS and the BSI-12.

Patient characteristics

At intake, age, gender, and education level (low: primary through lower secondary education/high: higher secondary education through university) were assessed. The Dutch version of the MINI International Neuropsychiatric Interview Plus (MINI-Plus, Sheehan *et al.*, 1998; Van Vliet, Leroy, & Van Megen, 2000) was used to collect DSM-IV-TR diagnostic information (type of anxiety disorder, the number of simultaneously occurring anxiety disorders, presence of a comorbid mood disorder, presence of a comorbid somatoform disorder, comorbid alcohol or substance abuse or dependence). The MINI-Plus has good psychometric properties, with inter-rater reliability between .88 and 1.00 and test–retest reliability between .76 and .93, and adequate validity compared to the Composite International Diagnostic Interview-1 (Lecrubier *et al.*, 1997).

Personality characteristics were assessed using the Dimensional Assessment of Personality Pathology short form (DAPP-SF, Van Kampen, De Beurs, & Andrea, 2008), the abbreviated version of the DAPP-BQ (Livesley, Jang, & Vernon, 1998). The DAPP-SF consists of 136 items on a 5-point Likert scale (1–5), which can be converted into subscales by taking the average of the subscale items. 18 subscales exist (i.e., submissiveness, cognitive distortion, identity problems, affective lability, stimulus seeking, compulsiveness, restricted expression, callousness, oppositionality, intimacy problems, rejection, anxiousness, conduct problems, suspiciousness, social avoidance, narcissism, insecure attachment, and self-harm). Higher scores are associated with pathology, whereas lower scores indicate normality. The DAPP-SF has good internal consistency with Cronbach's alphas ranging from .78 to .89 across subscales (Van Kampen *et al.*, 2008).

Analyses

Sample categorical characteristics are presented as number (percentage), and continuous variables are presented as mean (M) (\pm standard deviation [SD]). To describe overall concordance, the Pearson's product–moment correlation coefficient between the BAS and the BSI-12 was computed. To quantify discordance on individual patient level, Z -scores were computed. The minimal difference for the two discordant groups was set at 1 SD ; that is, a difference of at least 1 Z -score was categorized as discordant (Dorz *et al.*, 2004). This resulted in three levels of concordance: observer-rated (BAS) \approx self-report (BSI-12), observer-rated (BAS) < self-report (BSI-12), and observer-rated (BAS) > self-report (BSI-12). We compared the two discordant groups (observer-rated [BAS] < self-report [BSI-12] and observer-rated [BAS] > self-report [BSI-12]) with the concordant group (observer-rated [BAS] \approx self-report [BSI-12], reference group) with regard to age, gender, education level, anxiety diagnosis, the number of simultaneously occurring anxiety disorders, presence of comorbid mood-, somatoform- and alcohol- or substance-abuse or dependence disorders, and the subscales of the DAPP-SF using univariable multinomial logistic regression. Finally, in order to identify predictors of discordance, all variables that were significantly associated in univariable analyses were entered in a multivariable model. All analyses were two-tailed, significance was set at $p < .05$, and correction for multiple testing using Bonferroni was performed.

Although all patients in our sample were diagnosed with one or more anxiety disorders according to DSM-IV-TR criteria using a validated diagnostic instrument (MINI-Plus, Van Vliet *et al.*, 2000), we have no information regarding primary diagnosis or focus of treatment as patients in our cohort were referred by their general practitioner for the treatment of mood, anxiety, or somatoform disorder. Therefore, some patients in our sample, while meeting diagnostic criteria for an anxiety disorder, might not have been

seeking treatment primarily for an anxiety disorder (but instead for a mood or somatoform disorder). Therefore, as a sensitivity analysis, we repeated analyses with only those subjects who met criteria for at least moderate severity on either the BAS or the BSI-12 so as to guarantee substantial anxiety severity. Moderate severity was defined as $BAS \geq 10.38$ (Schat *et al.*, 2013; Tyrer *et al.*, 1984) and/or $BSI-12 \geq 6$ (Roy-Byrne *et al.*, 2010; Schat *et al.*, 2013). We used SPSS version 20.0 (IBM Corp., Armonk, NY, USA).

Results

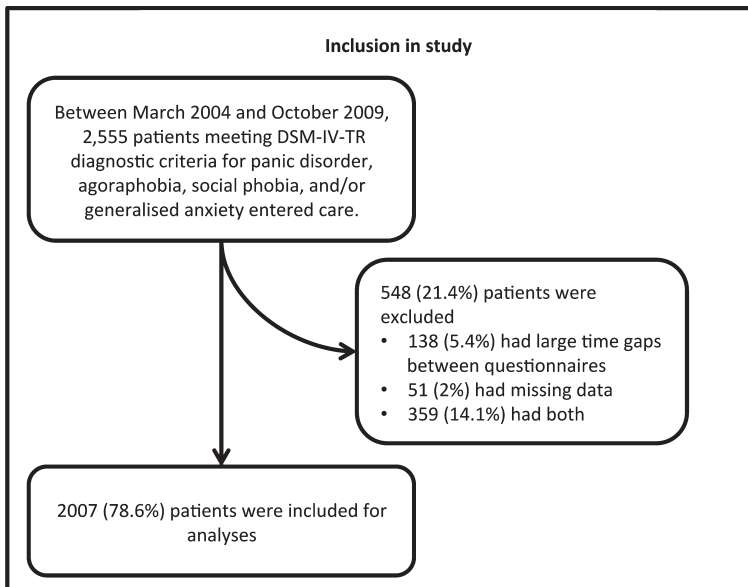
Sample characteristics

Between March 2004 and October 2009, a total of 2,555 patients met DSM-IV-TR diagnostic criteria for panic disorder with or without agoraphobia, agoraphobia without panic, social phobia, and generalized anxiety disorder. A total of 548 (21.4%) patients had to be excluded, as they had not completed all questionnaires or had a large time gap (of more than 21 days) between the completion of distinct questionnaires (see Table S1). In total, 138 (5.4%) patients were excluded solely due to a prolonged time gap, 359 (14.1%) patients had both a prolonged time gap and missing data, and 51 (2.0%) were excluded solely due to missing data. Therefore, 2,007 patients were included for analyses. Figure 1 shows a flow chart of inclusion and exclusion. Although differences between included and excluded patients were significant at $p < .05$ with regard to age, generalized anxiety disorder, comorbid depression, BSI-12 score, and DAPP-SF subscale cognitive distortion (see Table S2), these differences were very small, as indicated by eta-squared values that ranged from .002 to .005 and Cramer's phi's that ranged from .04 to .07. A total of 1282 (63.8%) patients were female and the mean age of the total sample was 36.0 years ($SD = 11.7$). Panic disorder with or without agoraphobia occurred in 784 (39.1%) patients, and agoraphobia without panic occurred in 430 (21.4%) patients; 650 (32.4%) patients were diagnosed with social phobia, and 430 (21.4%) had generalized anxiety disorder. A total of 577 (28.7%) patients met diagnostic criteria for multiple simultaneously occurring anxiety disorders. Comorbid mood disorder occurred in 908 (45.2%) patients, while 239 (11.9%) patients met diagnostic criteria for comorbid somatoform disorder.

Concordance

The relationship between observer-rated (BAS) and self-report (BSI-12) anxiety severity in the total sample was positive and strong; $r = .59$, $n = 2,007$, $p < .001$. Based on their standardized difference score, patients were categorized in three groups: observer-rated (BAS) = self-report (BSI-12), where $Z-BAS$ was equal to $Z-BSI-12 \pm 1$ ($n = 1,510$; 75.2%); observer-rated (BAS) < self-report (BSI-12), where $Z-BAS$ was equal to $Z-BSI-12$ minus at least 1 ($n = 244$; 12.2%); and observer-rated (BAS) > self-report (BSI-12), where $Z-BAS$ was equal to $Z-BSI-12$ plus at least 1 ($n = 253$; 12.6%). Table 1 shows sample characteristics for these three groups of outpatients. Table 2 shows associations of patient characteristics with level of concordance (BAS > BSI-12 and BAS < BSI-12 compared to BAS \approx BSI-12). Patients with higher observer-rated anxiety severity relative to self-report anxiety severity (BAS > BSI-12) did not differ from the concordant group (BAS \approx BSI-12) with regard to any of the patient characteristics.

Compared to patients whose anxiety severity scores were concordant (BAS \approx BSI-12), patients whose observed anxiety severity was lower than what they reported (BAS < BSI-



DSM-IV-TR denotes Diagnostic Statistical Manual-fourth edition, text revision

Figure 1. Flow chart of inclusion. DSM-IV-TR denotes Diagnostic Statistical Manual, Fourth Edition, text revision.

12) more often had a diagnosis of panic disorder with or without agoraphobia (OR = 2.06; 95% CI = 1.58–2.70; $p < .001$) and were less often diagnosed with social phobia (OR = 0.51; 95% CI = 0.37–0.71; $p < .001$). In addition, patients with lower observed compared to self-rated anxiety severity had higher scores on DAPP-SF subscales cognitive distortion (OR = 1.55; 95% CI = 1.36–1.77; $p < .001$), identity problems (OR = 1.45; 95% CI = 1.26–1.67; $p < .001$), affective lability (OR = 1.74; 95% CI = 1.47–2.05; $p < .001$), oppositionality (OR = 1.31; 95% CI = 1.13–1.53; $p = 0.001$), anxiousness (OR = 1.35; 95% CI = 1.15–1.58; $p < .001$), suspicion (OR = 1.32; 95% CI = 1.16–1.49; $p < .001$), and insecure attachment (OR = 1.31; 95% CI = 1.16–1.48; $p < .001$) (Table 2). Figure 2 shows ORs and CIs for significant associations. When, after checking for collinearity, these variables were entered in multivariable logistic regression, panic disorder with or without agoraphobia (OR = 1.73; 95% CI = 1.29–2.31; $p < .001$), social phobia (OR = 0.57; 95% CI = 0.40–0.81; $p = 0.002$), cognitive distortion (OR = 1.20; 95% CI = 1.00–1.45; $p = 0.05$), and affective lability (OR = 1.50 95% CI = 1.15–1.95; $p = .003$) best predicted lower observed compared to self-reported anxiety severity (Table 3). When analyses were repeated with only those patients who met criteria for at least moderate severity (BAS ≥ 10.38 (Tyrer *et al.*, 1984; Schat *et al.*, 2013)) and/or BSI-12 ≥ 6 (Roy-Byrne *et al.*, 2010; Schat *et al.*, 2013) ($n = 1,852$) results did not change (results not shown).

Discussion

We set out to describe concordance between an observer-rated (BAS) and self-report (BSI-12) measure of anxiety severity in a naturalistic sample of outpatients diagnosed with panic disorder with or without agoraphobia, agoraphobia without panic, social phobia,

Table 1. Baseline characteristics of 2,007 outpatients diagnosed with DSM-IV-TR anxiety disorders per level of concordance between observer-rated (BAS) and self-report (BSI-12) anxiety questionnaire

	Observer \approx self <i>n</i> = 1510	Observer > self <i>n</i> = 244	Observer < self <i>n</i> = 253
Age	36.12 (11.67)	36.91 (11.64)	34.64 (11.84)
Gender			
Male	549 (36.4%)	83 (34.0%)	93 (36.8%)
Female	961 (63.6%)	161 (66.0%)	160 (63.2%)
Education level			
High	912 (60.4%)	138 (55.6%)	135 (53.4%)
Low	598 (39.6%)	106 (43.4%)	118 (46.6%)
Panic disorder with or without agoraphobia	572 (37.9%)	71 (29.1%)	141 (55.7%)
Agoraphobia without panic	321 (21.3%)	53 (21.7%)	56 (22.1%)
Social phobia	498 (33.0%)	101 (41.4%)	51 (20.2%)
Generalized anxiety disorder	330 (21.9%)	60 (24.6%)	40 (15.8%)
Multiple anxiety disorders	409 (27.1%)	83 (34.0%)	85 (33.6%)
Comorbid mood	660 (43.7%)	130 (53.3%)	118 (46.6%)
Comorbid somatoform	170 (11.3%)	40 (16.4%)	29 (11.5%)
Comorbid alcohol	87 (5.8%)	14 (5.7%)	14 (5.5%)
Comorbid substance	71 (4.7%)	10 (4.1%)	11 (4.3%)
DAPP-SF subscales			
Submissiveness	3.07 (0.95)	3.05 (0.92)	3.09 (0.99)
Cognitive distortion	2.37 (0.97)	2.22 (0.91)	2.82 (1.04)
Identity problems	3.10 (1.02)	3.08 (0.94)	3.46 (0.97)
Affective lability	3.26 (0.87)	3.24 (0.76)	3.64 (0.84)
Stimulus seeking	2.08 (0.78)	2.01 (0.73)	2.19 (0.89)
Compulsiveness	2.92 (0.92)	2.91 (0.95)	3.11 (1.00)
Restricted expression	3.27 (0.85)	3.31 (0.84)	3.33 (0.85)
Callousness	1.78 (0.59)	1.73 (0.57)	1.90 (0.64)
Oppositionality	2.80 (0.88)	2.76 (0.84)	3.01 (0.89)
Intimacy problems	2.35 (0.80)	2.45 (0.87)	2.45 (0.86)
Rejection	2.30 (0.82)	2.14 (0.83)	2.36 (0.86)
Anxiousness	3.46 (0.92)	3.53 (0.84)	3.70 (0.91)
Conduct problems	1.43 (0.58)	1.39 (0.53)	1.49 (0.66)
Suspicion	2.23 (1.02)	2.30 (0.99)	2.53 (1.08)
Social avoidance	3.18 (1.07)	3.25 (1.06)	3.29 (1.08)
Narcissism	2.42 (0.82)	2.30 (0.80)	2.54 (0.84)
Insecure attachment	3.04 (1.11)	2.88 (1.05)	3.37 (1.15)
Self-harm	1.62 (0.90)	1.60 (0.79)	1.75 (0.96)

n denotes number of patients.

Note. Data are mean (standard deviation) or number (percentage) when appropriate. The MINI International Neuropsychiatric Interview Plus (MINI-Plus) was used to collect DSM-IV-TR diagnostic information (type of anxiety disorder, the number of simultaneously occurring anxiety disorders, presence of a comorbid mood disorder, presence of a comorbid somatoform disorder, comorbid alcohol or substance abuse or dependence); DSM-IV-TR denotes Diagnostic Statistical Manual, Fourth Edition, text revision; BAS denotes Brief Anxiety Scale; BSI-12 denotes Brief Symptom Inventory 12-item version; DAPP-SF denotes Dimensional Assessment of Personality Pathology short form; *n* denotes number of patients.

Table 2. Associations between patient characteristics and level of concordance between observer-rated (observer; BAS) and self-report (self; BSI-12) anxiety questionnaires in 2,007 outpatients with DSM-IV-TR anxiety disorders

Reference group: observer \approx self (<i>n</i> = 1,510)	Observer > self (<i>n</i> = 244)		Observer < self (<i>n</i> = 253)	
	OR (95% CI)	<i>p</i>	OR (95% CI)	<i>p</i>
Age	1.01 (0.99–1.02)	.33	0.99 (0.98–1.00)	.06
Female gender (vs. male)	1.11 (0.83–1.47)	.48	0.98 (0.75–1.30)	.90
Low education level (vs. high)	0.85 (0.65–1.12)	.26	0.75 (0.57–0.98)	.04
Panic disorder with or without agoraphobia	0.67 (0.50–0.90)	.009	2.06 (1.58–2.70)	<.001
Agoraphobia without panic	1.03 (0.74–1.43)	.87	1.05 (0.76–1.45)	.75
Social phobia	1.44 (1.09–1.89)	.01	0.51 (0.37–0.71)	<.001
Generalized anxiety disorder	1.17 (0.85–1.60)	.34	0.67 (0.47–0.96)	.03
Multiple anxiety disorders	0.72 (0.54–0.96)	.03	0.73 (0.55–0.98)	.03
Comorbid mood disorder	1.47 (1.12–1.93)	.005	1.13 (0.86–1.47)	.39
Comorbid somatoform disorder	1.55 (1.06–2.25)	.02	1.02 (0.67–1.55)	.92
Comorbid alcohol abuse/dependence	1.00 (0.56–1.78)	.99	0.96 (0.54–1.71)	.89
Comorbid substance abuse/dependence	0.87 (0.44–1.70)	.68	0.92 (0.48–1.76)	.92
DAPP-SF subscales				
Submissiveness	0.99 (0.86–1.14)	.85	1.03 (0.89–1.18)	.71
Cognitive distortion	0.85 (0.73–0.98)	.03	1.55 (1.36–1.77)	<.001
Identity problems	0.99 (0.86–1.13)	.82	1.45 (1.26–1.67)	<.001
Affective lability	0.98 (0.84–1.14)	.78	1.74 (1.47–2.05)	<.001
Stimulus seeking	0.90 (0.75–1.08)	.25	1.18 (1.01–1.39)	.04
Compulsiveness	0.98 (0.85–1.14)	.82	1.24 (1.07–1.43)	.003
Restricted expression	1.06 (0.91–1.25)	.46	1.09 (0.93–1.27)	.30
Callousness	0.85 (0.67–1.08)	.27	1.37 (1.11–1.69)	.004
Oppositionality	0.95 (0.82–1.11)	.54	1.31 (1.13–1.53)	<.001
Intimacy problems	1.16 (0.99–1.36)	.07	1.15 (0.98–1.36)	.07
Rejection	0.78 (0.66–0.93)	.005	1.10 (0.94–1.29)	.23
Anxiousness	1.09 (0.94–1.26)	.27	1.35 (1.15–1.58)	<.001
Conduct problems	0.87 (0.67–1.11)	.24	1.18 (0.95–1.45)	.13
Suspicion	1.07 (0.94–1.22)	.30	1.32 (1.16–1.49)	<.001
Social avoidance	1.07 (0.94–1.21)	.31	1.11 (0.97–1.25)	.12
Narcissism	0.84 (0.71–0.99)	.04	1.19 (1.01–1.39)	.04
Insecure attachment	0.88 (0.78–0.99)	.04	1.31 (1.16–1.48)	<.001
Self-harm	0.97 (0.83–1.13)	.66	1.16 (1.01–1.33)	.04

n denotes number of patients. Results statistically significant at significant at $p < .05$ after Bonferroni correction for multiple testing have been printed in bold.

Note. Data present odds ratios relative to the reference group 'no discordance' (Z-BAS % (Z-BSI '1) $n = 1,510$) obtained in univariable multinomial logistic regression. Bonferroni correction for multiple testing was made. The MINI International Neuropsychiatric Interview Plus (MINI-Plus) was used to collect DSM-IV-TR diagnostic information (type of anxiety disorder, the number of simultaneously occurring anxiety disorders, presence of a comorbid mood disorder, presence of a comorbid somatoform disorder, comorbid alcohol or substance abuse or dependence); DSM-IV-TR denotes Diagnostic Statistical Manual, Fourth Edition, text revision; BAS denotes Brief Anxiety Scale; BSI-12 denotes Brief Symptom Inventory 12-item version; OR denotes odds ratio; 95% CI denotes 95% confidence interval; DAPP-SF denotes Dimensional Assessment of Personality Pathology short form; *n* denotes number of patients. Results statistically significant at $p < .05$ after Bonferroni correction for multiple testing have been printed in bold.

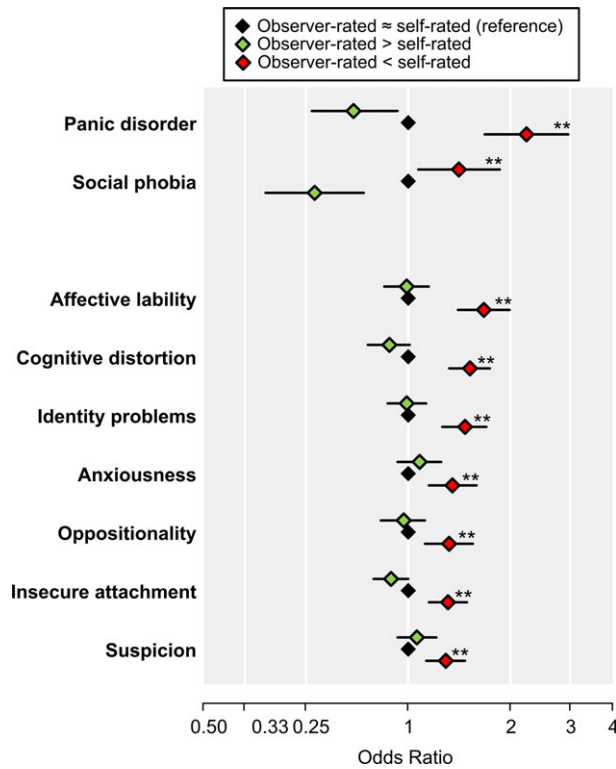


Figure 2. Significant associations of patient characteristics with discordance between observer-rated (BAS) and self-report (BSI-12) measures of anxiety severity, relative to concordant patients. Odds ratios with error bars representing 95% confidence intervals are shown for significant associations of patient characteristics in the observer-rated (BAS) < self-report (BSI-12) group relative to concordant patients; results for the observer-rated (BAS) > self-report (BSI-12) group are also shown; none of these were significant after Bonferroni correction; data were analysed using univariable multinomial logistic regression analysis; the MINI International Neuropsychiatric Interview Plus (MINI-Plus) was used to collect DSM-IV-TR diagnostic information; affective lability, cognitive distortion, identity problems, anxiousness, oppositionality, insecure attachment, and suspicion are all subscales of the Dimensional Assessment of Personality Pathology short form; ** denotes significant at $p < .05$ after Bonferroni correction for multiple testing; BAS denotes Brief Anxiety Scale; BSI-12 denotes Brief Symptom Inventory 12-item version. [Colour figure can be viewed at wileyonlinelibrary.com]

and/or generalized anxiety disorder. In addition to describing the level of concordance in our sample, we studied associations between patient characteristics and discordance, and examined which patient characteristics best predicted discordance.

The overall correlation between observer-rated and self-report measures of anxiety severity in our sample was positive and strong. However, for a substantial group of patients, considerable discordance existed, with observer-rated anxiety severity exceeding self-reported anxiety severity in 12.2% of patients, and lower observer-rated than self-reported anxiety in 12.6% of patients. These percentages were comparable to results from a previous study applying the same methodology with depressed patients and depression measures, in which these groups consisted of 17.7% and 15.5% of patients, respectively (Dorz *et al.*, 2004). Although a by definition unknown proportion of the discordance must be attributed to random measurement error, the discordance could be partially associated

Table 3. Predictors of higher self-reported compared to observer-rated anxiety severity in 2,007 outpatients with DSM-IV-TR anxiety disorders

Reference group observer \approx self ($n = 1,510$)	Observer < self ($n = 253$)	
	OR (95% CI)	p
Intercept -3.88		
Panic disorder with or without agoraphobia	1.73 (1.29–2.31)	<.001
Social phobia	0.57 (0.40–0.81)	.002
DAPP-SF subscales		
Cognitive distortion	1.20 (1.00–1.45)	.05
Identity problems	1.15 (0.92–1.45)	.23
Affective lability	1.50 (1.15–1.95)	.003
Oppositionality	0.99 (0.81–1.45)	.91
Anxiousness	0.87 (0.68–1.10)	.24
Suspicion	1.07 (0.90–1.26)	.46
Insecure attachment	1.02 (0.88–1.19)	.76

Note. Data present odds ratios relative to the reference group 'no discordance' ($Z\text{-BAS} \approx (Z\text{-BSI} \pm 1)$ $n = 1,510$) obtained in multivariable logistic regression. The MINI International Neuropsychiatric Interview Plus (MINI-Plus) was used to collect DSM-IV-TR diagnostic information (type of anxiety disorder, the number of simultaneously occurring anxiety disorders, presence of a comorbid mood disorder, presence of a comorbid somatoform disorder, comorbid alcohol or substance abuse or dependence); DSM-IV-TR denotes Diagnostic Statistical Manual, Fourth Edition, text revision; BAS denotes Brief Anxiety Scale; BSI-12 denotes Brief Symptom Inventory 12-item version; OR denotes odds ratio; 95% CI denotes 95% confidence interval; DAPP-SF denotes Dimensional Assessment of Personality Pathology short form.

with patient characteristics. The group of patients with higher observed anxiety relative to self-reported anxiety did not differ from the concordant group with regard to patient characteristics. The group of patients that had lower observed anxiety than what was reported, however, was more often diagnosed with panic disorder with or without agoraphobia compared to the concordant group, but less often had social phobia. Also, this group scored higher on personality aspects cognitive distortion, identity problems, affective lability, oppositionality, anxiousness, suspicion, and insecure attachment than the concordant group. Lower observed than self-reported anxiety severity was best predicted by panic disorder with or without agoraphobia, social phobia, cognitive distortion, and affective lability.

Our finding of higher prevalence of panic disorder with or without agoraphobia, and lower presence of social phobia in the group with lower observed relative to self-reported anxiety severity, may have several explanations. First, it may reflect a difference in response styles between patients in both diagnostic groups: Panic disorder is characterized by intense panic experiences, which may lead to high self-ratings of anxiety severity, while social phobia often entails feelings of shame and self-effacing, which may in turn result in underreporting of anxiety severity. Alternatively, it may be explained in terms of item content. Possibly, the BAS contains more items on (excessive or unreasonable) fear of social situations, whereas the BSI-12 may put more emphasis on the (more physical) symptoms of panic disorder. Inspection of the items of both instruments (Table 4), as well as the presence of an opposite (although not significant after Bonferroni correction) association in the $\text{BAS} > \text{BSI-12}$ group, supports this thought.

Table 4. Item content of the observer-rated Brief Anxiety Scale (BAS) and the self-report Brief Symptom Inventory 12-item version (BSI-12)

Items BAS (observer-rated)	Items BSI-12 (self-report)
Autonomic disturbances reported	Nervousness
Aches and pains	Faintness
Inner tension	Pains in chest
Hypochondriasis	Suddenly scared
Worrying over trifles	Feeling fearful
Phobias	Nausea
Hostile feelings	Trouble getting breath
Reduced sleep	Numbness
Autonomic disturbances observed	Feeling weak
Muscular tension	Feeling tense
	Spells of panic
	Feeling restless

The association between discordance and personality characteristics is in line with previous findings in depression. We found that patients whose observed anxiety severity was lower than their self-reported anxiety severity scored higher on cognitive distortion, identity problems, affective lability, oppositionality, anxiousness, suspicion, and insecure attachment than concordant patients. As higher DAPP-SF scores indicate elevated chances of personality pathology, these findings fit previous reports of higher prevalence of personality disorders in patients who reported higher depression severity relative to their observed depression severity (Dorz *et al.*, 2004; Rane *et al.*, 2010). Our findings are furthermore in agreement with previous more specific reports of high neuroticism, low extraversion, low agreeableness (Enns *et al.*, 2000), novelty seeking, and reward dependence (Carter *et al.*, 2010) in patients whose observed depression severity was lower than their self-reported depression severity. Lower observed anxiety severity compared to self-reported anxiety severity was best predicted by panic disorder with or without agoraphobia, social phobia, cognitive distortion, and affective lability. This finding indicates that especially in patients with positive scores on these characteristics, the use of both self-report and observer-rated instruments is merited. Together, these findings may indicate that personality pathology, especially in clusters B and C, adds to the suffering experienced by patients in a manner that is not readily observed or recognized by research nurses or caregivers.

We did not replicate previous findings of associations between level of concordance and age (Carter *et al.*, 2010; Dorz *et al.*, 2004; Enns *et al.*, 2000), gender (Carter *et al.*, 2010; Jolly *et al.*, 1994), education level (Enns *et al.*, 2000), and marital status (Dorz *et al.*, 2004). As these findings pertained largely to depression (Dorz *et al.*, 2004; Enns *et al.*, 2000), it is possible that they are specific to depression and depression instruments and do not generalize to anxiety. This thought is supported by reports by Jolly *et al.* (1994), who found that for adolescent boys with various psychiatric diagnoses, discordance on depression instruments was higher than for girls, whereas this difference did not exist for anxiety measures, although no such findings have been reported for age, education level, and marital status.

As previous research on concordance has mainly focussed on depression and no comparison between the scores on questionnaires used in this study has been made before, our results are novel. Other strong points are our large sample size, the use of trained research nurses, our naturalistic sample, and the thorough assessment of

personality pathology. However, several potential limitations exist. First of all, we have no information on the order in which observer-rated and self-report instruments were administered; possibly, the order in which instruments were administered varied, which in turn might have influenced the level of concordance (Jolly *et al.*, 1994). Second, we had no information on primary diagnosis, which might have been associated with discordance. Third, although we excluded patients with large time gaps (more than 21 days) between the administration of the BAS and the BSI-12, smaller time gaps were allowed and incidentally occurred, which may have influenced the results. However, although this might have influenced the number of cases in which discordance occurred, it is unlikely to have influenced the association between patient characteristics and discordance. Fourth, we used two different instruments to measure anxiety severity; a comparison between scores on an observer-rated and self-report version of the same instrument would have been preferable as this would have ruled out the explanation of associations through item content. Although our results are in line with previous findings in mood disorders, they do not allow attribution of findings to the nature of the instruments used instead of differences in item content. Future studies should be aimed at replication using different measures of anxiety severity or, ideally, a self-report and observer-rated version of a single instrument. Furthermore, the DAPP-SF, which was used to assess personality pathology, is a self-report instrument. It must be noted therefore that our results reflect that those who have a tendency to score high on self-reported relative to observed anxiety severity tend to score high on self-reported cognitive distortion, identity problems, affective lability, oppositionality, anxiousness, suspicion, and insecure attachment. Although the vast majority of measures of personality characteristics are self-report (Conelly & Deniz, 2010), replication of this study with an observer-rated measure of personality could provide valuable new insights. In addition, we have no information on rater characteristics, which might be associated with discordance as well (Carter *et al.*, 2010; Van Noorden *et al.*, 2010). On a related topic, although we did not find associations between patient characteristics and higher observer-rated compared to self-reported anxiety severity, this does not imply the discordance in this group necessarily resulted entirely from random measurement error. Possibly, this type of discordance is related to rater characteristics, or patient characteristics that were not measured in this study. Finally, as our data are cross-sectional, our findings do not allow for causal interpretation regarding contributions of personality characteristics to discordance.

Our results demonstrate that using a single instrument in the assessment of patients' anxiety severity could give rise to a one-sided view of pathology. Self-reported anxiety and observer-rated anxiety severity are not identical constructs. Even if a self-report version and an observer-rated version of the same instrument had been used, these would not have measured a singular construct. Whereas observer-rated anxiety severity holds a ranking of anxiety severity in the context of the clinical expertise of and impression by the professional who does the rating, self-reported anxiety severity represents a subjective report of degree of suffering, the level to which an individual feels they can endure this, and a self-reflective capacity. Therefore, although in most patients a high degree of concordance is to be expected, objective ranking and subjective degree of suffering should not be expected to match in each single patient. Although both measures provide an estimate of anxiety severity within their respective context, they are often used interchangeably. Our results demonstrate that this could result in the failure to note clinically relevant information. Clinicians and researchers should be aware of the distinction between both measurement methods as results show associations between assessment method and diagnosis and personality pathology. Our findings have practical

implications for research as well as practice. As in psychiatry, patients' subjective experiences are central, it is relevant to note that for a substantial group of patients, self-reported anxiety severity does not match observed anxiety severity. In clinical studies, observer-rated instruments, while generally regarded as the primary source of information, may not suffice when measuring anxiety severity and change in anxiety severity, as (changes in) subjectively experienced anxiety severity may go unnoted. In clinical practice, on the other hand, the growing reliance on cheaper self-report scales may obscure anxiety severity in a substantial group of patients for whom anxiety severity would be rated higher by a trained research nurse. Also, our results demonstrate that for those patients who rate their anxiety as more severe than a trained observer would, personality pathology might be a complicating factor. Possibly, patients who have high cognitive distortion, identity problems, affective lability, oppositionality, anxiousness, suspicion, and insecure attachment are less willing or capable to express their symptom severity to an observer. Finally, the potential presence of personality pathology is highly relevant for treatment as personality pathology may require a different approach and special attention. Therefore, in conclusion, we argue that a multimethod approach to psychiatric assessment, consisting of self-report as well as observer-rated instruments, disorder-specific as well as generic instruments, and covering aspects of symptoms, personality, and psychosocial functioning, although more expensive, is highly preferable to the use of a single self-report instrument in baseline assessment of psychopathology (Enns *et al.*, 2000; Moller, 2000).

Acknowledgements

We would like to thank the patients who participated in this study and the research nurses.

References

- Asberg, M., Montgomery, S. A., Perris, C., Schalling, D., & Sedvall, G. (1978). Comprehensive psychopathological rating-scale. *Acta Psychiatrica Scandinavica*, Supplement, 271, 5–27. <https://doi.org/10.1111/j.1600-0447.1978.tb02357.x>
- Carter, J. D., Frampton, C. M., Mulder, R. T., Luty, S. E., & Joyce, P. R. (2010). The relationship of demographic, clinical, cognitive and personality variables to the discrepancy between self and clinician rated depression. *Journal of Affective Disorders*, 124, 202–206. <https://doi.org/10.1016/j.jad.2009.11.011>
- Conelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, 136, 1092–1122.
- De Beurs, E., Den Hollander-Gijsman, M. E., Van Rood, Y. R., Van der Wee, N. J. A., Giltay, E. J., Van Noorden, M. S., ... Zitman, F. G. (2011). Routine outcome monitoring in the Netherlands: Practical experiences with a web-based strategy for the assessment of treatment outcome in clinical practice. *Clinical Psychology & Psychotherapy*, 18, 1–12. <https://doi.org/10.1002/cpp.696>
- De Beurs, E., & Zitman, F. G. (2006). De Brief Symptom Inventory (BSI) De betrouwbaarheid van een handzaam alternatief voor de SCL-90. *Maandblad Geestelijke Volksgezondheid*, 61, 120–141.
- Dorz, S., Borgherini, G., Conforti, D., Scarso, C., & Magni, G. (2004). Comparison of self-rated and clinician-rated measures of depressive symptoms: A naturalistic study. *Psychology & Psychotherapy: Theory, Research and Practice*, 77, 353–361. <https://doi.org/10.1348/1476083041839349>
- Dunlop, B. W., Li, T., Kornstein, S. G., Friedman, E. S., Rothschild, A. J., Pedersen, R., Ninan, P., ... Trivedi, M. H. (2011). Concordance between clinician and patient ratings as predictors of

- response, remission, and recurrence in major depressive disorder. *Journal of Psychiatric Research*, *45*, 96–103. <https://doi.org/10.1016/j.jpsychires.2010.04.032>
- Enns, M. W., Larsen, D. K., & Cox, B. J. (2000). Discrepancies between self and observer ratings of depression. The relationship to demographic, clinical and personality variables. *Journal of Affective Disorders*, *60*, 33–41. [https://doi.org/10.1016/S0165-0327\(99\)00156-1](https://doi.org/10.1016/S0165-0327(99)00156-1)
- Franke, G. H., Ankerhold, A., Haase, M., Jager, S., Togel, C., Ulrich, C., & Frommer, J. (2011). The usefulness of the brief symptom inventory 18 (BSI-18) in psychotherapeutic patients. *Psychotherapie Psychosomatik Medizinische Psychologie*, *61*, 82–86. <https://doi.org/10.1055/s-0030-1270518>
- Goekoop, J. G., Knoppertvanderklein, E. A. M., Hoeksema, T., Klinkhamer, R. A., Vangaalen, H. A. E., & Vandervelde, E. A. (1991). The interrater reliability of a dutch version of the comprehensive psychopathological rating-scale. *Acta Psychiatrica Scandinavica*, *83*, 202–205. <https://doi.org/10.1111/j.1600-0447.1991.tb05525.x>
- Hamilton, M. (1976). Comparative value of rating scales. *British Journal of Clinical Pharmacology*, *3*, 58–60. <https://doi.org/10.1111/j.1365-2125.1976.tb03714.x>
- Hopko, D. R., Bourland, S. L., Stanley, M. A., Beck, J. G., Novy, D. M., Averill, P. M., & Swann, A. C. (2000). Generalized anxiety disorder in older adults: Examining the relation between clinician severity ratings and patient self-report measures. *Depression & Anxiety*, *12*, 217–225. [https://doi.org/10.1002/1520-6394\(2000\)12:4%3C217:AID-DA5%3E3.0.CO;2-6](https://doi.org/10.1002/1520-6394(2000)12:4%3C217:AID-DA5%3E3.0.CO;2-6)
- Jolly, J. B., Wiesner, D. C., Wherry, J. N., Jolly, J. M., & Dykman, R. A. (1994). Gender and the comparison of self and observer ratings of anxiety and depression in adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *33*, 1284–1288. <https://doi.org/10.1097/00004583-199411000-00009>
- Lecrubier, Y., Sheehan, D. V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K. H., . . . Dunbar, G. C. (1997). The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: Reliability and validity according to the CIDI. *European Psychiatry*, *12*, 224–231. [https://doi.org/10.1016/S0924-9338\(97\)83296-8](https://doi.org/10.1016/S0924-9338(97)83296-8)
- Livesley, W. J., Jang, K. L., & Vernon, P. A. (1998). Phenotypic and genetic structure of traits delineating personality disorder. *Archives of General Psychiatry*, *55*, 941–948. <https://doi.org/10.1001/archpsyc.55.10.941>
- Lubaczewski, S., Shepherd, J., Fayyad, R., & Guico-Pabia, C. J. (2014). Real-world disparities between patient- and clinician-reported outcomes: Results from a disease-specific program in depression and anxiety. *Professional Case Management*, *19*, 63–74. <https://doi.org/10.1097/NCM.0000000000000015>
- Moller, H. J. (2000). Rating depressed patients: Observer- vs self-assessment. *European Psychiatry*, *15*, 160–172. [https://doi.org/10.1016/S0924-9338\(00\)00229-7](https://doi.org/10.1016/S0924-9338(00)00229-7)
- Rane, L. J., Fekadu, A., Wooderson, S., Poon, L., Markopoulou, K., & Cleare, A. J. (2010). Discrepancy between subjective and objective severity in treatment-resistant depression: Prediction of treatment outcome. *Journal of Psychiatric Research*, *44*, 1082–1087. <https://doi.org/10.1016/j.jpsychires.2010.03.020>
- Roy-Byrne, P., Craske, M. G., Sullivan, G., Rose, R. D., Edlund, M. J., Lang, A. J., . . . Stein, M. B. (2010). Delivery of evidence-based treatment for multiple anxiety disorders in primary care: A randomized controlled trial. *Journal of the American Medical Association*, *303*, 1921–1928. <https://doi.org/10.1001/jama.2010.608>
- Schat, A., Van Noorden, M. S., Noom, M. J., Giltay, E. J., Van der Wee, N. J., Vermeiren, R. R. J. M., & Zitman, F. G. (2013). Predictors of outcome in outpatients with anxiety disorders: The Leiden routine outcome monitoring study. *Journal of Psychiatric Research*, *47*, 1876–1885. <https://doi.org/10.1016/j.jpsychires.2013.09.005>
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., . . . Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (MINI): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, *59*, 22–33.

- Tyrer, P., Owen, R. T., & Cicchetti, D. V. (1984). The brief scale for anxiety – A subdivision of the comprehensive psychopathological rating-scale. *Journal of Neurology Neurosurgery and Psychiatry*, *47*, 970–975. <https://doi.org/10.1136/jnnp.47.9.970>
- Van Kampen, D., De Beurs, E., & Andrea, H. (2008). A short form of the Dimensional Assessment of Personality Pathology-Basic Questionnaire (DAPP-BQ): The DAPP-SF. *Psychiatry Research*, *160*, 115–128. <https://doi.org/10.1016/j.psychres.2007.05.002>
- Van Noorden, M. S., Giltay, E. J., Den Hollander-Gijsman, M. E., Van der Wee, N. J. A., Van Veen, T., & Zitman, F. G. (2010). Gender differences in clinical characteristics in a naturalistic sample of depressive outpatients: The Leiden Routine Outcome Monitoring Study. *Journal of Affective Disorders*, *125*, 116–123. <https://doi.org/10.1016/j.jad.2009.12.007>
- Van Vliet, I. M., Leroy, H., & Van Megen, H. J. G. M. (2000). MINI internationaal neuropsychiatrisch interview nederlandse versie 5.0.0, 5 ed. Utrecht.
- Zabora, J., BrintzenhofeSzoc, K., Jacobsen, P., Curbow, B., Piantadosi, S., Hooker, C., . . . Derogatis, L. (2001). A new psychosocial screening instrument for use with cancer patients. *Psychosomatics*, *42*, 241–246. <https://doi.org/10.1176/appi.psy.42.3.241>

Received 13 November 2015; revised version received 4 May 2017

Supporting Information

The following supporting information may be found in the online edition of the article:

Table S1. Breakdown of reasons for exclusion from analyses in 2,555 outpatients diagnosed with DSM-IV-TR anxiety disorders.

Table S2. Comparison of included and excluded cases in 2,555 outpatients diagnosed with DSM-IV-TR anxiety disorders categorical variables.