Chapter 2

Predicting outcomes of mood, anxiety and somatoform disorders: The Leiden routine outcome monitoring study

Martijn van Noorden*
Esther van Fenema*
Nic van der Wee
Yanda van Rood
Ingrid Carlier
Frans Zitman
Erik Giltay

**Joint first authors*

J Affect Disord. 2012 Dec 15;142(1-3):122-31

Abstract

Background: Mood, anxiety and somatoform (MAS) disorders are highly prevalent and comorbid disorders with substantial mutual comorbidity and a large disease burden. Cross-diagnostic predictors for poor outcome of MAS disorders in routine clinical practice are lacking. The aim of this study was to predict outcomes in outpatients with MAS disorders using Routine Outcome Monitoring (ROM) data.

Methods: A cohort study in, 892 adult MAS patients in a naturalistic outpatient psychiatric specialty care setting, with a replication cohort of 1392 patients. Poor outcome was defined as a <50% reduction (compared to baseline) on the self-report Brief Symptom Inventory (BSI) *or* a score of ≥3 on the observer-rated Clinical Global Improvement Scale (CGI), during follow-up during up to 2 years.

Results: In multivariable Cox regression models, independent and replicated predictors for poor outcome were higher age, having comorbid MAS disorders or a somatoform disorder, dysfunctional personality traits (i.e., tendency to self-harm, intimacy problems, affective lability), and a low reported general health status.

Conclusions: In routine clinical care, specific patient profiles may need special interventions to minimize the risk of poor outcome.

Introduction

Mood, anxiety and somatoform (MAS) disorders consistently rank as a highly prevalent group of disorders, responsible for a considerable burden of disease as measured by several indicators. [1] Patients suffering from single MAS disorders often display poor outcomes (with high disability, long duration of the illness, and high risk of recurrence) with more than half of all patients not achieving remission after first-line treatment, while the presence of one or more comorbid MAS disorders contributes to this even further. [2,2-5] Longitudinal studies on patients suffering from single MAS disorders have identified several predictors for adverse outcomes. In general, poorer remission rates from depressive disorders were independently predicted by being unmarried [6-9], being unemployed [8,9], a lower level of education [10], a greater severity of depressive symptoms [6,9,11,12], concomitant symptoms of pain, comorbid anxiety disorders [7,11-13], and borderline personality disorder. [12,14] Poorer remission rates from anxiety disorders were predicted by a higher severity of the anxiety symptoms, concomitant symptoms of pain, comorbid depression [13] and prevalent personality disorders. [15,16] Predictors of low remission rates for somatoform disorders were a lower level of education [10], concomitant symptoms of depression [17], a greater severity of the somatoform symptoms [18], and suffering from a comorbid personality disorder. [19]

However, the available studies did not include patients with different MAS diagnoses concomitantly and showed important methodological differences. Sample size ranging from 165 [6] to 1996 [14], a duration of follow-up ranging from 3 months [9] to 5 years [16], the use of samples from highly selected populations [8,9], a great variety in the number of determinants included in multivariable models, and not taking comorbidity into account.

Given the high frequency of mutual comorbidity in MAS disorders in routine clinical practice, cross-diagnostic predictors of poor outcomes would have clinical relevance. The present study aims to identify possible cross-diagnostic predictors of poor outcomes in a naturalistic cohort of 892 outpatients with MAS disorders during up to 2 years of follow-up, and to replicate the results in a second, independent cohort of 1392 MAS outpatients.

Methods

Routine outcome monitoring (ROM)

In our routine outcome monitoring infrastructure, all outpatients referred to Rivierduinen (RD) or Leiden University Medical Center (LUMC) for treatment of a MAS disorder are routinely assessed with an extensive psychometric battery at baseline during intake and, if treatment is initiated, repeatedly every 3–4 months during treatment. [20,21] Patients with non-MAS disorders (e.g., psychotic disorders, main diagnosis of personality disorder) are referred to other specialised care settings within our institutions. In ROM, data on diagnosis and complaint severity are collected systematically to assess treatment effectiveness in everyday clinical practice. In our setting, ROM is performed by trained psychiatric research nurses who are not involved in treatment. A group-wise quality control and calibration among research nurses ensures quality maintenance during data collection.[20] All questionnaires are completed on touch-screen computers, to prevent missing data within questionnaires. Patients with insufficient mastery of the Dutch language are ineligible. During the study period, on average 80% of the referred patients were assessed with ROM. ROM data are primarily used for diagnosis and to inform clinicians and patients about treatment progress. Patient-identifiable data are removed from the database in order to secure patients' confidentiality. The use of these anonymised data for research purposes has been approved by the Ethical Review Board of the LUMC.

During the first ROM session, a standardized diagnostic interview is administered as well as observer-rated and self-report scales, both generic and disorder-specific. In addition, demographic variables are collected (an overview of instruments is available at http://www.lumc.nl/psychiatry/ROM-instruments). At intake, current and lifetime Axis-I diagnoses according to the diagnostic and statistical manual of mental disorders, fourth edition, text revision (DSM-IV-TR) were established using the mini-international neuropsychiatric interview-plus. [22] The MINI-Plus has good psychometric properties, with inter-rater reliability ranging from 0.88 to 1.00, test-retest reliability ranging from 0.76 to 0.93 and adequate validity compared to the composite international diagnostic interview. [23]

Assessment of outcomes

For the purpose of the present study, we used data collected with two generic, disorder-independent instruments: the patient-rated brief symptom inventory (BSI) and the observer-rated clinical global impression scale (CGI). The BSI is a 53-item version of the symptom check list [24] that assesses psychopathological symptoms in eight symptom domains on a 5-point Likert scale (from 0 'not at all' through 4 'extremely'). The BSI subscales are: somatisation, obsessive compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic fear, paranoia and psychoticism. The total score is computed by taking the mean score of all individual items (range 0–4). The BSI has shown good internal consistency, with Cronbach's alpha ranging from 0.71 to 0.85 [20,24] and test-retest reliability coefficients ranging from 0.68 to 0.91. The convergent validity has proven to be very good. [24] The CGI is a simple standardized observer-rated assessment tool for making global assessments of the severity of illness. [25] The main item 'severity of illness', measured on a 7-point Likert scale (from 1 'normal, not at all ill' through 7 'among the most extremely ill patients'), was used in the present analyses (CGI-S). The CGI is widely used in medical care and clinical research as an outcome measure because of its face validity and practicability. [26]

Within the cross-diagnostic design of our study, poor outcome was defined according to the rather stringent criteria of a <50% reduction of the baseline BSI score [27] **or** a CGI-S score of ≥3 during a maximum of 2 years of follow-up. Only when a patient reached a BSI reduction >50% **and** CGI-S score 1 or 2 during follow up, this patient was considered a responder at that time point. When there was a discontinuation of follow-up measurements, the survival time was censored at the last time point at which a ROM assessment was completed when response was not achieved.

Patients

We used two cohorts of adult outpatients who were referred to RD and LUMC for treatment of a MAS disorder between 2004 and 2009. The first (initial) cohort consisted of 2876 patients with a ROM-baseline assessment (aged 18–65 years) with one or more current DSM-IV-TR MAS disorders according to the MINI-Plus, included between 1 January 2004 and 31 December 2006. The second (replication) cohort consisted of 2966 patients (aged 18–65 years) with DSM-IV-TR MAS disorders according to the MINI-Plus with a ROM-baseline

assessment included between 1 January 2007 and 31 December 2009. Patients with a lifetime bipolar disorder or primary psychotic illness were excluded, but a diagnosis of major depression with psychotic features according to the MINI-Plus was allowed.

We then first excluded all patients with low baseline severity, defined as a BSI total score of <0.5 or a score of <3 on the CGI-S, because these patients with minimal severity at baseline were unlikely to receive treatment and follow-up assessments. [27] Then we excluded all patients who did not have follow-up ROM assessments after the ROM baseline assessment. Finally, patients with incomplete data (missings on one or more outcome variables or predictor variables) were excluded.

In the initial cohort, 892 of 2876 patients (31.0%) were included, and in the replication cohort 1392 of 2966 patients (46.9%). There were no differences in gender between the included/excluded patients in either of the cohorts. In the initial cohort, no difference in age existed between the included/excluded patients. Included patients more often had a higher education than excluded patients (59.0% vs. 42.7%, χ^2 (2)=25,243, **p**<0.001. In the replication cohort, excluded patients were slightly younger (mean age at baseline assessment 37.53 years vs. 38.53 years, **t**(2964)=2.16, **P**=0.03) than participating patients. As expected based on the exclusion criteria, excluded patients had lower mean BSI and CGI-S scores than included patients in both the initial cohort (mean BSI score 1.29 vs. 1.40, **t**(2786)=3.84, **p**<0.001 and mean CGI-S score 4.04 vs. 4.14, **t**(2441)=2.67, **P**=0.008, respectively), and the replication cohort (1.23 vs. 1.42, **t**(2963)=7.39, **p**<0.001 and 4.02 vs. 4.17, **t**(2964)=4.86, **p**<0.001, respectively).

Psychiatrists and clinical psychologists or psychotherapists in the LUMC and RD provided treatment according to the principle of stepped-care and based on the Dutch evidence-based treatment guidelines, consisting mainly of pharmacotherapy, psychotherapy, or a combination of both. [28] Treatment and therapist characteristics were not taken into account in the present analyses, because baseline treatment-independent patient characteristics were the focus of the present analysis.

Potential predictor variables

Categorical predictor variables

An extensive set of clinical and demographic variables was available. Demographic variables were obtained at baseline with a self-report

questionnaire. A Dutch ethnic background was assumed when the patient and both parents were born in the Netherlands. Marital status was categorized in 'married' (which also included living together in a relationship), 'divorced or widowed', and 'never married'. Housing situation was categorized in 'living alone', 'living with partner', and 'living with family'. Lower education was defined as having completed elementary school or lower general secondary education. Employment situation was categorized in 'working full-time', 'working part-time', 'retired/unemployed', and 'on sick leave'.

DSM-IV-TR diagnostic categories (as established with the MINI-Plus) were categorized as: mood disorders (major depressive disorder or dysthymic disorder), anxiety disorders (panic disorder, social phobia, generalized anxiety disorder, posttraumatic stress disorder or obsessive-compulsive disorder) and somatoform disorders (hypochondriasis, pain disorder, body dysmorphic disorder, somatisation disorder or undifferentiated somatoform disorder), alcohol abuse or dependence, or drug abuse or dependence. If more than 1 MAS diagnosis was established with the MINI-Plus, the patients was assumed to have MAS comorbidity.

Continuous predictor variables

Disorder-independent clinical variables were assessed with the following self-report and observer-rated scales:

The observer-rated abbreviated comprehensive psychopathological rating scale (CPRS) consists of the Montgomery–Äsberg depression rating scale (MADRS), the Brief Anxiety Scale (BAS) and a scale that assesses psychomotor inhibition. (INH) [29] The MADRS has an internal consistency (Cronbach's alpha) of 0.86, and an inter-rater reliability coefficient of 0.65–0.97. [30]

The global assessment of functioning (GAF) scale is a rating scale for evaluating 'psychological, social and occupational functioning on a hypothetical continuum of mental illness' (Axis V of the DSM-IV-TR). The GAF score is measured on a scale of 0–100, and the most severe condition on any of the three dimensions provides the overall score. The GAF score has a modest reliability, which strongly depends on the training and calibration of the raters. [31]

All participants also completed the mood and anxiety symptoms questionnaire (MASQ), a questionnaire that measures the dimensions of the tripartite model of anxiety and depression. [32] The MASQ is a 90-item self-report questionnaire

that consists of five symptom dimensions and has good psychometric properties with Cronbach's alphas ranging from 0.91 to 0.96 across subscales. [32]

The Dimensional Assessment of Personality Pathology-Short Form (DAPP-SF) was administered to assess maladaptive personality traits. [33] The DAPP-SF is the Short version of the DAPP-BQ, a self-report scale that consists of 18 subscales with a total of 136 items. Scores are on a 5-point Likert scale (1–5), and scores of subscales are computed by taking the sum scores of the subscale items. The DAPP-SF has good psychometric properties with Cronbach's alphas ranging from 0.78 to –0.89 across subscales. [33]

Generic health status was assessed with the short form-36 (SF-36), a 36-item self-report questionnaire that measures health status in eight domains: physical functioning, social functioning, physical problems, emotional problems, mental health, vitality, bodily pain and general health. [34] The SF-36 has good psychometric properties with Cronbach's alphas ranging from 0.66 to 0.94 across subscales. [34]

Statistical analysis

Using descriptive statistics, baseline characteristics are described as number (percentage) or mean (standard deviation [SD]), when appropriate. Univariable hazard ratios (HR) of poor outcome were computed according to baseline categorical and continuous predictor variables in the initial cohort. HRs were calculated for poor outcome based on the BSI and the CGI-S separately, as well as on the more conservative combined CGI-S *or* BSI criterion for poor outcome. To allow comparison of the obtained effect sizes on different predictor variables, standardised *z*-scores were calculated (as the difference between measured values and mean, divided by the SD). Since higher scores on the GAF scale and SF-36 subscales correspond with better functioning and health, we subtracted original scores from 100, and used these inverted scores in the analyses to facilitate the comparability among HRs.

All predictor variables from the initial cohort that had yielded HRs with a *p*-value <0.10 in univariable analyses were subsequently selected for an initial forward stepwise multivariable Cox regression model. The criteria used for both selection and removal were 0.10. Independent predictors of poor outcome as previously defined in the initial cohort were used for replication in the independent replication cohort. These predictors were selected in the

replication cohort and forced in a multivariable Cox regression model. Because exact dates of reaching the defined endpoint of treatment response were unknown, a sensitivity analysis was performed in which we re-analysed the date of the ROM measurement during which response was achieved, taking the mid time point between the last two measurements. This mid-time point is probably more close to the 'true' time point at which response was reached. Additional sensitivity analyses were performed with less conservative response criteria had to be met (i.e., CGI-S *and* BSI criterion for poor outcome), and with less stringent exclusion criteria (low baseline severity defined as baseline BSI score <0.5 *or* CGI-S score of 1 or 2). Moreover, we performed a sensitivity analysis in which all patients with a single somatoform disorder were excluded from the analyses. All further tests were two-tailed with a $p<0.05$ denoting statistical significance. The software used was SPSS version 17.0 (SPSS Inc., Chicago, Ill, USA).

Results

Sample and demographic characteristics

Table 1 shows the predictor variables and DSM-IV-TR diagnoses at baseline of 892 and 1392 subjects of the initial and replication cohorts, respectively. A total of 62.4% of the initial cohort and 64.6% of the replication cohort were female. The mean age in the initial cohort was 38.3 (SD 11.6) years and was 38.5 (SD 12.5) years in the replication cohort. A Dutch ethnic background was found in 81.2% in the initial cohort and in 84.3% in the replication cohort. In both cohorts, the majority of patients was living with a partner, and more than half of the patients were not (currently) working. A single mood disorder was the most prevalent disorder in both cohorts (33.4% and 29.3%, respectively), followed by a single anxiety disorder (22.9% and 26.9%, respectively) and a single somatoform disorder (5.9% and 4.7%, respectively). MAS comorbidity was present in 37.8% and 39.0% of the patients in the initial and replication cohorts, respectively. Due to the naturalistic character of the study, we expected relatively large attrition rates. Roughly every 6 months, the total amount of patients with follow-up ROM assessments decreased with 50%.

Univariable predictors of poor outcome

Table 2 shows the univariable categorical correlates of poor outcome in the initial cohort. Lower education was associated with a higher HR of poor outcome, which was consistent among the two outcome variables. Using the 'combined BSI or CGI-S poor outcome' variable, lower education resulted in a 41% higher chance of poor outcome. Employment status was also consistently associated with poor outcome, with retired subjects, unemployed subjects and subjects being on sick leave having a higher chance of poor outcome on any of the three outcome variables. Suffering from MAS comorbidity predicted a 51% higher HR of poor outcome using the 'combined BSI **or** CGI-S poor outcome' variable, as compared with having a single mood disorder. Patients suffering from a single somatoform disorder three times more chance of poor outcome using the 'combined BSI **or** CGI-S poor outcome' variable, as compared with having a single mood disorder.

Table 3 presents the univariable continuous correlates of poor outcome in the initial cohort. Again, overall the findings were largely consistent when using either the BSI or CGI-S or the combined poor outcome criterion. An increase of 1 SD in age corresponded with a 22% higher chance of poor outcome using the 'combined BSI **or** CGI-S poor outcome' endpoint. The largest and most consistent effect sizes on the two and the combined outcome measures were found for the DAPP-SF subscales cognitive distortion, identity problems, affective lability, intimacy problems and self-harm, and for the SF-36 subscales poor physical functioning and poor general health.

Independent predictors of poor outcome

Table 4 shows the multivariable HRs of poor outcome on the 'combined BSI **or** CGI-S poor outcome' variable in both the initial and replication cohorts.

When using the group with a single mood disorder as the reference group, we found that having a single anxiety disorder was not significantly associated with a higher chance of poor outcome. However, the relatively small groups (of 53 and 66 subjects) with a single somatoform disorder had consistently more than three times higher chance of poor outcome. The group with MAS DSM-IV-TR comorbidity also had a 52% higher chance of poor outcome in both cohorts, as compared with the single mood disorder group.

Next, a 1-SD increase in the DAPP-SF affective lability subscale was associated with a 22% and 36% higher chance of poor outcome in the initial and replication

cohorts, respectively. A 1-SD increase in the DAPP-SF Intimacy Problems subscale resulted in a 28% higher chance of poor outcome in the initial cohort, and 16% in the replication cohort. Another independent predictor of poor outcome was the self-harm subscale of the DAPP-SF, with 1 SD increase of scores resulting in 16% and 11% higher chance of poor outcome, respectively. Finally, one subscale from the SF-36 was strongly and independently associated with treatment response in both cohorts, i.e., unfavorable scores on the General Health subscale that predicted 22% and 26% higher chances of poor outcome, respectively Fig. 1.

Fig. 2 shows the six independent predictors of poor outcome in subjects from the combined initial and replication cohorts in Kaplan–Meier curves, representing the cumulative incidence of naturalistic treatment response with a maximum follow-up time of 2 years. In all sensitivity analyses (as described in the statistical analyses section), the same independent predictors were found (data not shown).

Discussion

In this naturalistic cohort study in 892 outpatients with MAS disorders and an independent replication cohort of 1392 outpatients, we found that having an older age, MAS comorbidity, a somatoform disorder, high scores on the personality dimensions affective lability, intimacy problems and self-harm, and a poor general health were independently associated with poor outcome after up to 24 months of follow-up. Our samples were broadly representative of outpatients with major depressive, anxiety and somatoform disorders (or a combination of these disorders) treated in a naturalistic psychiatric secondary care setting. Detailed assessment of patient characteristics before the start of treatment proved to be useful in predicting poor outcomes.

Our data confirms and adds to the growing body of evidence about risk factors for poor outcome in MAS disorders. First, in previous studies older age predicted slower recovery among outpatients with depression [8] and outpatients with depression and/or anxiety. [13] Physical illness, disability, social isolation and loneliness may become more common with advancing age, which may adversely affect treatment outcome of comorbid psychiatric disorders. [35] Second, comorbidity has repeatedly been related to poor outcomes, [3,4,9,11-

13,17] which may be due to the higher burden of disease and severity. [3,4] Third, somatic symptoms and somatic comorbidity have been linked to poor outcomes, [10,36] but longitudinal studies in patients with somatoform disorders are scarce. Fourth, adverse personality traits were found to interfere with and compromise therapy in many epidemiological studies. [12,15,16,19] Personality disorders at baseline in a sample of 303 patients with a depressive disorder were associated with long-term adverse outcomes. [37] In a study among 514 patients with generalized anxiety disorder, social phobia and panic disorder, the presence of a personality disorder adversely influenced the time to remission during 5 years of follow-up. [16] Finally, somatic conditions and poor general health have often been associated with depressive and anxiety disorders. [7,10,38,39] Besides a direct association, co-medication (e.g., opiate analgesics, antihypertensives, corticosteroids and interferons), may further adversely have induced or detrimented MAS disorders.

Taking our results together, a MAS patient who meets the profile of being elderly, suffering from comorbid MAS disorders or a somatoform disorder, with cluster B personality traits and a reported poor general health, is at risk of poor outcome or chronicity in clinical outpatient practice. Our findings stress the importance of assessing the strengths and vulnerabilities of the patient's personality, as high levels of affective lability, tendency to self-harm and intimacy problems, indicative of poor coping styles and dysfunctional interpersonal behavior, predicted poor outcome and should be a focus of attention during treatment. [33] MAS patients with more pronounced personality traits may benefit from specific therapeutic approaches, such as targeting cluster B personality characteristics, or cognitive behavioral therapy or dialectical behavior therapy that combines cognitive-behavioral techniques with concepts of distress tolerance, acceptance, and mindful awareness. [40] Self-perceived poor general health and somatoform disorders also deserve more research on focused treatment. Although patients suffering from MAS disorders are likely to make illness attributions with somatic symptoms and experience illness behavior, any underlying medical conditions should be clarified and treated if necessary, especially in the elderly. Also, these patients might benefit from specific interventions to improve their physical condition, e.g., consulting a physiotherapist or dietician, and collaborative care management. [39] Screening and assessment instruments can be routinely applied at baseline to systematically detect psychiatric comorbidity, adverse

personality characteristics and self-perceived poor general health.

Our study has several strengths. First, this was a cross-diagnostic study in naturalistic samples with high comorbidity rates, and with broad inclusion and few exclusion criteria representative of a 'real world' setting. Second, we confirmed our findings in a large independent replication cohort. Third, a structured clinical interview was used to diagnose psychiatric disorders, and detailed self-report and observer-rated information was available on individual characteristics including personality traits. Fourth, although no rating scale has generally been accepted as the gold standard to assess social and functional aspects of recovery in MAS patients, measures such as the BSI and the CGI are clinically useful instruments to evaluate treatment outcomes across the diverse MAS disorders and to assess the patient population in its entirety. [41,42] Finally, all assessments were made by trained research nurses who were not involved in treatment.

Our study also has several potential limitations. First, we expected and found substantial attrition rates, as many patients remitted from their MAS disorder and completed only one follow-up ROM assessment. A considerable loss to follow-up was also common in other studies with a 'real world' approach such as the STAR*D trial, that reached a loss to follow-up of 30% after step II of the study. [43] Unfortunately, we were unable to report the reasons of attrition (e.g., remission and discharge, no-show, aborted treatment). Second, there is a lack of information on the specific treatments given, thus treatment could not be taken into account in our analyses. We assume that patients were treated according to evidence-based guidelines; however, as our study represents a 'real world' setting, guideline concordance is often less strict than in a controlled setting. Nevertheless, a previous analysis in our ROM cohort showed that treatment broadly followed guidelines, and consisted of psychotherapy, pharmacotherapy, and combined therapy. [28,44] Third, no information on psychiatric history, family history and somatic comorbidity of the patients was available for our cohorts. Fourth, we used the DAPP-SF for assessment of personality traits and identifying patients more likely to suffer from DSM-IV-TR personality disorders; however, the Axis I MAS disorder may have confounded and precluded a valid assessment of personality disorders. Fifth, the BSI may not be specific enough to fully capture clinical changes in all MAS disorders, especially in somatoform disorders. [45] Finally, the use of Cox regression analyses implies that subjects are censored upon reaching the

defined endpoint of treatment response; therefore, the information about a possible relapse after an initial response could not be taken into account.
In summary, we discovered and replicated cross-diagnostic predictors that identify outpatients with MAS disorders who are at risk for poor outcome in a naturalistic outpatient treatment setting. These predictors are: an older age, diagnosis of a somatoform disorder or MAS comorbidity, affective lability, intimacy problems, self-harm, and a poorer general health. Since our patient cohorts were representative a large naturalistic treatment-seeking population in outpatient psychiatric specialty care, our findings contribute to the existing literature on predictors for outcome in MAS disorders. As MAS disorders are highly prevalent and often invalidating conditions, future studies should focus on the identification of the most effective treatment modalities for the most vulnerable outpatient subgroups among the heterogeneous outpatient population with MAS disorders.

Role of funding source
No external funding was provided for this study.
Conflict of interest
All authors declare that they have no conflicts of interest.
Acknowledgment
The authors thank the research nurses for their help in data collection. There are no disclosures for any author. MS van Noorden and EM van Fenema had full access to all data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Wittchen HU, Jacobi F: Size and burden of mental disorders in Europe--a critical review and appraisal of 27 studies. Eur Neuropsychopharmacol 15:357-376, 2005.
2. Gaynes BN: Identifying difficult-to-treat depression: differential diagnosis, subtypes, and comorbidities. J Clin Psychiatry 70 Suppl 6:10-15, 2009.
3. Maier W, Falkai P: The epidemiology of comorbidity between depression, anxiety disorders and somatic diseases. Int Clin Psychopharmacol 14 Suppl 2:S1-S6, 1999.
4. de Waal MW, Arnold IA, Eekhof JA, et al: Somatoform disorders in general practice: prevalence, functional impairment and comorbidity with anxiety and depressive disorders. Br J Psychiatry 184:470-476, 2004.
5. Kessler RC, Berglund P, Demler O, et al: Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 62:593-602, 2005.
6. Meyers BS, Sirey JA, Bruce M, et al: Predictors of early recovery from major depression among persons admitted to community-based clinics: an observational study. Arch Gen Psychiatry 59:729-735, 2002.
7. Weinberger MI, Sirey JA, Bruce ML, et al: Predictors of major depression six months after admission for outpatient treatment. Psychiatr Serv 59:1211-1215, 2008.
8. Fournier JC, DeRubeis RJ, Shelton RC, et al: Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression. J Consult Clin Psychol 77:775-787, 2009.
9. Frank E, Cassano GB, Rucci P, et al: Predictors and moderators of time to remission of major depression with interpersonal psychotherapy and SSRI pharmacotherapy. Psychol Med 41:151-162, 2011.
10. McKenzie M, Clarke DM, McKenzie DP, et al: Which factors predict the persistence of DSM-IV depression, anxiety, and somatoform disorders in the medically ill three months post hospital discharge? J Psychosom Res 68:21-28, 2010.
11. Melartin TK, Rytsala HJ, Leskela US, et al: Severity and comorbidity predict episode duration and recurrence of DSM-IV major depressive disorder. J Clin Psychiatry 65:810-819, 2004.
12. Souery D, Oswald P, Massat I, et al: Clinical factors associated with treatment resistance in major depressive disorder: results from a European multicenter study. J Clin Psychiatry 68:1062-1070, 2007.
13. Penninx BW, Nolen WA, Lamers F, et al: Two-year course of depressive and anxiety disorders: Results from the Netherlands Study of Depression and Anxiety (NESDA). J Affect Disord 2011.
14. Skodol AE, Grilo CM, Keyes KM, et al: Relationship of Personality Disorders to the Course of Major Depressive Disorder in a Nationally Representative Sample. Am J Psychiatry 168:257-264, 2011.
15. Yonkers KA, Dyck IR, Warshaw M, et al: Factors predicting the clinical course of generalised anxiety disorder. Br J Psychiatry 176:544-549, 2000.
16. Massion AO, Dyck IR, Shea MT, et al: Personality disorders and time to remission in generalized anxiety disorder, social phobia, and panic disorder. Arch Gen Psychiatry 59:434-440, 2002.
17. Leiknes KA, Finset A, Moum T, et al: Course and predictors of medically unexplained pain symptoms in the general population. J Psychosom Res 62:119-128, 2007.
18. Hiller W, Leibbrand R, Rief W, et al: Predictors of course and outcome in hypochondriasis after cognitive-behavioral treatment. Psychother Psychosom 71:318-325, 2002.
19. Phillips KA, Pagano ME, Menard W, et al: Predictors of remission from body dysmorphic disorder: a prospective study. J Nerv Ment Dis 193:564-567, 2005.
20. de Beurs E, den Hollander-Gijsman ME, van Rood YR, et al: Routine outcome monitoring in the Netherlands: practical experiences with a web-based strategy for the assessment of treatment outcome in clinical practice. Clin Psychol Psychother 18:1-12, 2011.
21. van Noorden MS, Giltay EJ, den Hollander-Gijsman ME, et al: Gender differences in clinical characteristics in a naturalistic sample of depressive outpatients: the Leiden Routine Outcome Monitoring Study 5. J Affect Disord 125:116-123, 2010.

22. Sheehan DV, Lecrubier Y, Sheehan KH, et al: 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, 1998.
23. Lecrubier Y: The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. European psychiatry 12:224, 1997.
24. Derogatis LR, Melisaratos N: The Brief Symptom Inventory: an introductory report. Psychol Med 13:595-605, 1983.
25. Guy W. ECDEU Assessment Manual for Psychopharmacology. Rockville, MD, National Institute of Mental Health. 1976. Ref Type: Generic
26. Kadouri A, Corruble E, Falissard B: The improved Clinical Global Impression Scale (iCGI): development and validation in depression. BMC Psychiatry 7:7, 2007.
27. Roy-Byrne P, Craske MG, Sullivan G, et al: Delivery of evidence-based treatment for multiple anxiety disorders in primary care: a randomized controlled trial. JAMA 303:1921-1928, 2010.
28. Fenema van EM, van der Wee NJ, Bauer M, et al: Assessing adherence to guidelines for common mental disorders in routine clinical practice. Int J Qual Health Care 24:72-79, 2012.
29. Goekoop JG, Hoeksema T, Knoppert-Van der Klein EA, et al: Multidimensional ordering of psychopathology. A factor-analytic study using the Comprehensive Psychopathological Rating Scale. Acta Psychiatr Scand 86:306-312, 1992.
30. Montgomery SA, Asberg M: A new depression scale designed to be sensitive to change. Br J Psychiatry 134:382-389, 1979.
31. Vatnaland T, Vatnaland J, Friis S, et al: Are GAF scores reliable in routine clinical use? Acta Psychiatr Scand 115:326-330, 2007.
32. Watson D, Weber K, Assenheimer JS, et al: Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. J Abnorm Psychol 104:3-14, 1995.
33. van Kampen D, de BE, Andrea H: A short form of the Dimensional Assessment of Personality Pathology-Basic Questionnaire (DAPP-BQ): the DAPP-SF. Psychiatry Res 160:115-128, 2008.
34. Ware JE, Jr., Sherbourne CD: The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 30:473-483, 1992.
35. Luanaigh CO, Lawlor BA: Loneliness and the health of older people. Int J Geriatr Psychiatry 23:1213-1221, 2008.
36. Speckens AE, van Hemert AM, Bolk JH, et al: Unexplained physical symptoms: outcome, utilization of medical care and associated factors. Psychol Med 26:745-752, 1996.
37. Grilo CM, Stout RL, Markowitz JC, et al: Personality disorders predict relapse after remission from an episode of major depressive disorder: a 6-year prospective study. J Clin Psychiatry 71:1629-1635, 2010.
38. Mezuk B, Eaton WW, Albrecht S, et al: Depression and type 2 diabetes over the lifespan: a meta-analysis. Diabetes Care 31:2383-2390, 2008.
39. Katon WJ, Lin EH, Von KM, et al: Collaborative care for patients with depression and chronic illnesses. N Engl J Med 363:2611-2620, 2010.
40. Robins CJ, Chapman AL: Dialectical behavior therapy: current status, recent developments, and future directions. J Pers Disord 18:73-89, 2004.
41. Trivedi MH: Sensitizing clinicians and patients to the social and functional aspects of remission. J Clin Psychiatry 62 Suppl 19:32-35, 2001.
42. Klinkman MS: Assessing functional outcomes in clinical practice. Am J Manag Care 15:S335-S342, 2009.
43. Rush AJ, Trivedi MH, Wisniewski SR, et al: Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report Am J Psychiatry 163:1905-1917, 2006.
44. van der Lem R, van der Wee NJ, van VT, et al: The generalizability of antidepressant efficacy trials to routine psychiatric out-patient practice. Psychol Med 41:1353-1363, 2011.
45. Arnold LM, Leon T, Whalen E, et al: Relationships among pain and depressive and anxiety symptoms in clinical trials of pregabalin in fibromyalgia. Psychosomatics 51:489-497, 2010.

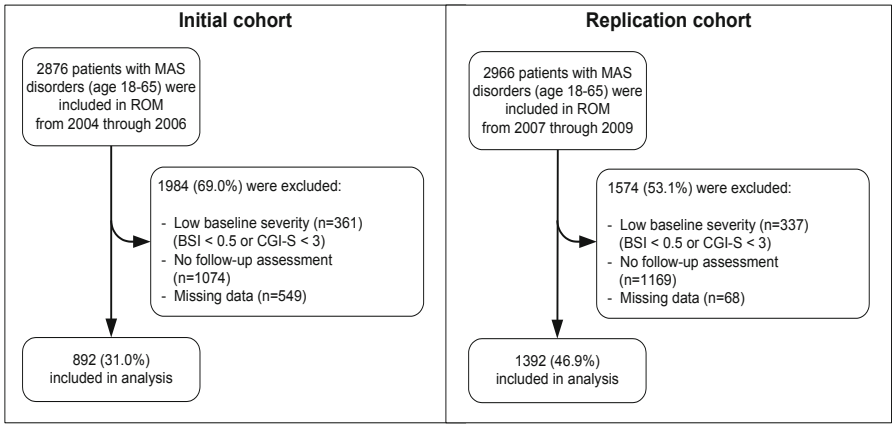

Figure 1. **Flowchart of the patients in the initial and replication cohorts.**

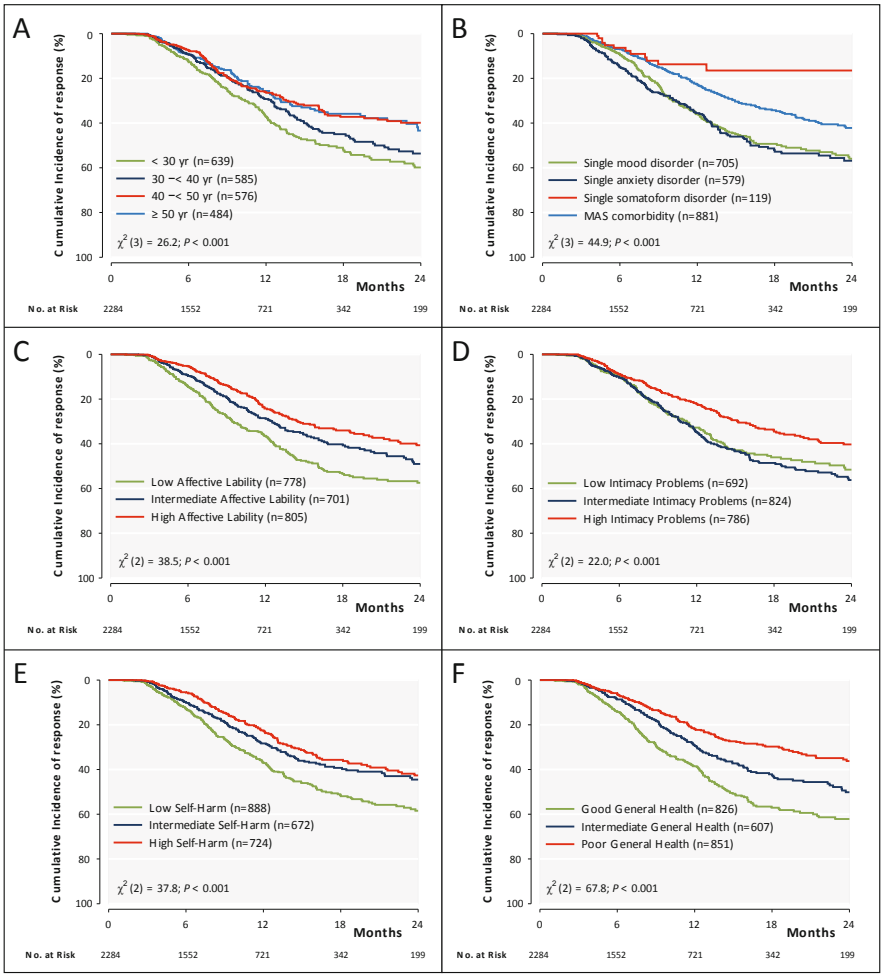

Figure 2. Kaplan-Meier curves for poor outcome according to baseline characteristics.

Kaplan-Meier curves are shown for the incidence of naturalistic treatment response based on a >50% improvement on the BSI *and* a score of 1 or 2 on the CGI-S among 2284 patients with a MAS disorder from the initial and replication cohorts combined. Patients were compared for age per decade (Box A), DSM-IV-TR diagnostic category (Box B), Affective Lability (Box C), Intimacy Problems (Box D), Self-Harm (Box E), and General Health (Box F). To facilitate graphical presentation of data, categorization of Affective Lability, Intimacy Problems, Self-Harm, and General Health into tertiles was applied.

Table 1. Baseline characteristics in 892 and 1392 outpatients with MAS disorders.

	Initial cohort (n=892)		Replication cohort (n=1392)	
Categorical predictor variables	n	%	n	%
Female gender	557	62.4	899	64.6
Dutch ethnic background	724	81.2	1173	84.3
Marital status				
- Married	454	50.9	763	54.8
- Divorced/ widowed	144	16.1	181	13.0
- Never married	294	33.0	449	32.2
Housing situation				
- Living alone	219	24.6	268	19.3
- Living with partner	472	52.9	772	55.4
- Living with family	201	22.5	353	25.3
Educational status				
- Lower education	366	41.0	568	40.7
- Higher education	526	59.0	825	59.3
Employment situation				
- Working full-time	196	22.0	289	20.8
- Working part time	199	22.3	332	23.9
- Retired/unemployed	232	26.0	363	26.1
- On sick leave	265	29.7	409	29.4
DSM-IV-TR diagnostic categories				
- Single Mood Disorder	298	33.4	408	29.3
- Single Anxiety Disorder	204	22.9	375	26.9
- Single Somatoform Disorder	53	5.9	66	4.7
- MAS Comorbidity	337	37.8	543	39.0
Alcohol abuse or dependence	48	5.4	74	5.3
Drug abuse or dependence	30	3.4	56	4.0
Continuous predictor variables	Mean (±SD)	IQR	Mean (±SD)	IQR
Age (years) at interview	38.3 (11.6)	28-47	38.5 (12.5)	28-49
BSI total score	1.4 (0.6)	0.9-1.8	1.4 (0.6)	0.9-1.8
CGI-Sscore	4.1 (0.8)	4.0-5.0	4.2 (0.8)	4.0-5.0
MADRS score	21.7 (8.7)	15-28	19.7 (8.4)	13-26
BAS score	16.5 (6.4)	12-21	15.9 (5.9)	12-20
GAF score	58.3 (8.2)	53-62	58.2 (6.5)	55-61

MAS denotes mood, anxiety and somatoform, SD denotes standard deviation, IQR denotes interquartile range, BSI denotes Brief Symptom Inventory, CGI–S denotes Clinical Global Impression-Severity, MADRS denotes Montgomery-Äsberg Depression Rating Scale, MAS denotes mood, anxiety and somatoform, BAS denotes Brief Anxiety Scale, GAF denotes Global Assessment of Functioning.

Table 2. Univariable hazard ratios of poor outcome according to baseline categorical predictor variables in 892 outpatients from the initial cohort.

	BSI poor outcome		CGI-S poor outcome		BSI or CGI-S poor outcome	
Categorical predictor variables	HR (95% CI)	P-value	HR (95% CI)	P-value	HR (95% CI)	P-value
Gender						
- Female	1.00		1.00		1.00	
- Male	1.15 (0.93-1.41)	0.20	1.07 (0.83-1.38)	0.62	1.09 (0.83-1.44)	0.52
Ethnic background						
- Dutch	1.00		1.00		1.00	
- Other	1.12 (0.87-1.44)	0.39	1.15 (0.83-1.59)	0.41	1.05 (0.75-1.47)	0.79
Marital status						
- Married	1.00		1.00		1.00	
- Divorced / widowed	1.18 (0.88-1.56)	0.27	1.38 (0.94-2.02)	0.10	1.35 (0.90-2.03)	0.15
- Never married	1.03 (0.83-1.28)	0.80	0.99 (0.76-1.3)	0.95	1.00 (0.75-1.34)	0.98
Housing situation						
- Living with partner	1.00		1.00		1.00	
- Living with family	1.01 (0.79-1.29)	0.95	0.89 (0.65-1.23)	0.49	0.94 (0.67-1.31)	0.70
- Living alone	1.12 (0.84-1.49)	0.45	0.95 (0.66-1.37)	0.79	1.07 (0.72-1.59)	0.73
Educational status						
- Higher education	1.00		1.00		1.00	
- Lower education	1.28 (1.05-1.57)	0.02	1.33 (1.02-1.72)	0.03	1.41 (1.06-1.87)	0.02
Employment situation						
- Working full-time	1.00		1.00		1.00	
- Working part time	1.42 (1.08-1.87)	0.01	1.22 (0.87-1.71)	0.25	1.32 (0.92-1.91)	0.14
- Retired/unemployed	1.80 (1.36-2.38)	<0.001	1.72 (1.22-2.43)	0.002	1.89 (1.30-2.77)	0.001
- On sick leave	1.78 (1.36-2.31)	<0.001	1.92 (1.37-2.70)	<0.001	1.83 (1.28-2.62)	0.001

Table 2. continued

	BSI poor outcome		CGI-S poor outcome		BSI or CGI-S poor outcome	
DSM-IV-TR diagnostic categories						
- Single Mood Disorder	Ref.		Ref.		Ref.	
- Single Anxiety Disorder	0.87 (0.67-1.13)	0.30	0.73 (0.54-0.98)	0.04	0.81 (0.59-1.12)	0.21
- Single Somatoform Disorder	1.84 (1.08-3.13)	0.03	2.06 (1.07-3.95)	0.03	2.95 (1.29-6.76)	0.01
- MAS Comorbidity	1.10 (0.87-1.38)	0.43	1.60 (1.18-2.17)	0.003	1.51 (1.09-2.09)	0.01
Alcohol abuse or dependence	1.38 (0.84-2.28)	0.20	1.79 (0.88-3.61)	0.11	1.71 (0.81-3.62)	0.16
Drug abuse or dependence	1.30 (0.71-2.37)	0.39	1.07 (0.55-2.09)	0.93	1.01 (0.5-2.05)	0.97

BSI denotes Brief Symptom Inventory, CGI-S denotes Clinical Global Impression-Severity, CI denotes Confidence Interval, HR denotes Hazard Ratio, Ref. denotes reference category. MAS denotes mood, anxiety and somatoform. Poor outcome on the BSI was defined as < 50% reduction of the baseline score, poor outcome on the CGI was defined as a score of ≥ 3 on the 7-point Likert scale after up to 24 months follow-up.

Table 3. Univariable hazard ratios of poor outcome according to baseline continuous predictor variables in 892 outpatients from the initial cohort.

	BSI poor outcome		CGI-S poor outcome		BSI or CGI-S poor outcome	
Continuous predictor variables	HR (95% CI)	P-value	HR (95% CI)	P-value	HR (95% CI)	P-value
Age	1.15 (1.04-1.27)	0.007	1.18 (1.04-1.34)	0.01	1.22 (1.06-1.41)	0.005
MADRS	1.10 (1.00-1.21)	0.06	1.31 (1.16-1.49)	<0.001	1.23 (1.08-1.4)	0.002
BAS	1.09 (0.99-1.21)	0.08	1.22 (1.07-1.38)	0.003	1.18 (1.03-1.36)	0.02
Low GAF score	1.10 (1.00-1.21)	0.04	1.26 (1.12-1.42)	<0.001	1.20 (1.06-1.36)	0.005
MASQ						
- Anhedonic Depression	1.10 (1.01-1.21)	0.04	1.31 (1.17-1.46)	<0.001	1.25 (1.10-1.41)	<0.001
- Anxious Arousal	1.10 (0.99-1.21)	0.08	1.22 (1.07-1.39)	0.003	1.17 (1.01-1.34)	0.03
- General Distress Depression	1.10 (1.00-1.21)	0.06	1.38 (1.22-1.57)	<0.001	1.30 (1.13-1.49)	<0.001
- General Distress Anxiety	1.07 (0.97-1.19)	0.16	1.23 (1.08-1.4)	0.001	1.15 (1.01-1.32)	0.04
- General Distress Mixed	1.10 (1.00-1.21)	0.05	1.32 (1.16-1.49)	<0.001	1.23 (1.08-1.41)	0.002

Table 3. continued

	BSI poor outcome		CGI-S poor outcome		BSI or CGI-S poor outcome	
Continuous predictor variables	HR (95% CI)	P-value	HR (95% CI)	P-value	HR (95% CI)	P-value
DAPP-SF						
- Submissiveness	1.01 (0.91-1.11)	0.90	1.02 (0.90-1.15)	0.77	0.99 (0.87-1.14)	0.93
- Cognitive Distortion	1.20 (1.08-1.33)	0.001	1.31 (1.15-1.5)	<0.001	1.26 (1.10-1.45)	0.001
- Identity Problems	1.14 (1.04-1.26)	0.006	1.31 (1.16-1.47)	<0.001	1.26 (1.11-1.43)	<0.001
- Affective Lability	1.11 (1.01-1.23)	0.03	1.27 (1.13-1.43)	<0.001	1.24 (1.08-1.41)	0.001
- Stimulus Seeking	1.03 (0.93-1.13)	0.61	1.01 (0.89-1.15)	0.85	1.01 (0.88-1.15)	0.91
- Compulsivity	1.08 (0.98-1.19)	0.13	1.10 (0.97-1.24)	0.14	1.09 (0.96-1.25)	0.20
- Restricted Expression	1.17 (1.06-1.30)	0.002	1.23 (1.09-1.40)	0.001	1.20 (1.05-1.38)	0.008
- Callousness	1.10 (0.99-1.22)	0.07	1.25 (1.11-1.42)	0.16	1.20 (1.06-1.38)	0.22
- Oppositionality	1.20 (1.09-1.32)	<0.001	1.25 (1.11-1.42)	<0.001	1.20 (1.06-1.38)	0.006
- Intimacy Problems	1.22 (1.10-1.34)	<0.001	1.25 (1.10-1.41)	0.001	1.28 (1.12-1.47)	<0.001
- Rejection	1.04 (0.94-1.15)	0.46	0.99 (0.88-1.12)	0.88	1.01 (0.88-1.15)	0.91
- Anxiousness	1.07 (0.97-1.18)	0.17	1.18 (1.05-1.33)	0.007	1.18 (1.03-1.34)	0.01
- Conduct Problems	1.18 (1.06-1.32)	0.004	1.16 (1.01-1.34)	0.03	1.21 (1.04-1.42)	0.02
- Suspiciousness	1.14 (1.03-1.26)	0.01	1.21 (1.06-1.38)	0.004	1.20 (1.05-1.39)	0.009
- Social Avoidance	1.13 (1.02-1.25)	0.02	1.27 (1.12-1.44)	<0.001	1.22 (1.06-1.39)	0.004
- Narcissism	0.97 (0.88-1.07)	0.60	0.92 (0.81-1.04)	0.18	0.92 (0.81-1.05)	0.24
- Insecure Attachment	0.99 (0.90-1.09)	0.87	1.12 (0.99-1.28)	0.07	1.02 (0.89-1.17)	0.76
- Self-Harm	1.18 (1.06-1.31)	0.002	1.34 (1.16-1.54)	<0.001	1.30 (1.12-1.51)	0.001

Table 3. continued

	BSI poor outcome		CGI-S poor outcome		BSI or CGI-S poor outcome	
Continuous predictor variables	HR (95% CI)	P-value	HR (95% CI)	P-value	HR (95% CI)	P-value
SF-36						
- Poor Physical Functioning	1.37 (1.23-1.53)	<0.001	1.57 (1.36-1.82)	<0.001	1.65 (1.40-1.94)	<0.001
- Poor Social Functioning	1.05 (0.95-1.16)	0.32	1.21 (1.07-1.36)	0.003	1.17 (1.02-1.33)	0.02
- Physical Problems	1.16 (1.05-1.28)	0.003	1.23 (1.10-1.39)	0.001	1.22 (1.07-1.39)	0.003
- Emotional Problems	1.06 (0.96-1.16)	0.27	1.07 (0.95-1.21)	0.25	1.03 (0.90-1.18)	0.62
- Poor Mental Health	1.04 (0.95-1.15)	0.41	1.27 (1.13-1.43)	<0.001	1.20 (1.05-1.36)	0.006
- Low Vitality	1.15 (1.04-1.27)	0.006	1.34 (1.18-1.52)	<0.001	1.29 (1.13-1.48)	<0.001
- Bodily Pain	1.21 (1.09-1.33)	<0.001	1.27 (1.12-1.45)	<0.001	1.28 (1.11-1.47)	<0.001
- Poor General Health	1.36 (1.23-1.49)	<0.001	1.51 (1.34-1.7)	<0.001	1.53 (1.34-1.74)	<0.001

BSI denotes Brief Symptom Inventory, CGI-S denotes Clinical Global Impression-Severity, CI denotes Confidence Interval, HR denotes Hazard Ratio, MADRS denotes Montgomery Äsberg Depression Rating Scale, BAS denotes Brief Anxiety Scale, GAF denotes Global Assessment of Functioning, MAS denotes mood, anxiety and somatoform, MASQ denotes Mood and Anxiety Symptoms Questionnaire, DAPP-SF denotes Dimensional Assessment of Personality Pathology- short form, SF-36 denotes Short Form 36. Poor outcome on the BSI was defined as < 50% reduction of the baseline score, poor outcome on the CGI-S was defined as a score of ≥ 3 on the 7-point Likert scale after up to 24 months follow-up. For both GAF and SF-36 inverted scores were computed by subtracting the original scores from 100 in order for higher scores to indicate poorer states. To increase comparability of HRs, Z-scores of original scores were used in Cox regression analyses.

Table 4. Multivariable hazard ratios of poor outcome in 892 and 1392 outpatients with MAS disorders.

	Initial cohort (n=892)		Replication cohort (n=1392)	
Variables	HR (95% CI)	P-value	HR (95% CI)	P-value
Age at interview (per decade)	1.14 (1-1.3)	0.05	1.15 (1.06-1.26)	0.001
Gender	0.87 (0.65-1.15)	0.31	0.93 (0.75-1.16)	0.53
DSM-IV-TR diagnostic categories				
- Single mood disorder	1.00 (Ref.)		1.00 (Ref.)	
- Single anxiety disorder	1.15 (0.81-1.63)	0.44	1.29 (0.99-1.67)	0.06
- Single somatoform disorder	3.24 (1.39-7.52)	0.006	3.17 (1.39-7.3)	0.01
- MAS comorbidity	1.52 (1.09-2.11)	0.01	1.52 (1.19-1.95)	0.001
DAPP-SF Affective Lability	1.22 (1.06-1.41)	0.007	1.36 (1.22-1.51)	<0.001
DAPP-SF Intimacy Problems	1.28 (1.11-1.48)	0.001	1.16 (1.04-1.29)	0.01
DAPP-SF Self-Harm	1.16 (0.98-1.36)	0.08	1.11 (0.99-1.25)	0.08
SF-36 Poor Physical Functioning	1.29 (1.06-1.56)	0.01	0.94 (0.83-1.06)	0.31
SF-36 Poor General Health	1.22 (1.04-1.43)	0.01	1.26 (1.12-1.42)	<0.001

CI denotes Confidence Interval, HR denotes Hazard Ratio, Ref. denotes reference group, MAS denotes mood, anxiety and somatoform, DAPP-SF denotes Dimensional Assessment of Personality Pathology short-form, SF-36 denotes Short Form 36. Poor outcome was defined as <50% reduction of the baseline Brief Symptom Inventory (BSI) score ***or*** a Clinical Global Impression-severity (CGI-S) score of ≥3 on the 7-point Likert-scale after up to 24 months follow-up. Independent predictors of poor outcome in the discovery cohort were identified using a forward stepwise model with gender and age at interview forced in the model; these predictors were forced in a model for replication.