

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/19950> holds various files of this Leiden University dissertation.

Author: Noorden, Martijn Sander van

Title: On real-world patients and real-world outcomes : the Leiden Routine Outcome Monitoring Study in patients with mood, anxiety and somatoform disorders

Issue Date: 2012-10-11

On real-world patients and real-world outcomes

The Leiden Routine Outcome Monitoring Study
in patients with mood, anxiety and somatoform disorders

Martijn Sander van Noorden

Martijn van Noorden

On real-world patients and real-world outcomes. The Leiden Routine Outcome Monitoring Study in patients with mood, anxiety and somatoform disorders

Print: GVO drukkers & vormgevers B.V. | Ponsen & Looijen, Ede
The Netherlands

© 2012 M.S. van Noorden, Leiden, The Netherlands

On real-world patients and real-world outcomes

The Leiden Routine Outcome Monitoring Study
in patients with mood, anxiety and somatoform disorders

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. P.F. van der Heijden,
volgens het besluit van het college voor Promoties
te verdedigen op 11 oktober 2012
klokke 16.15 uur

door

Martijn Sander van Noorden
geboren te Amsterdam in 1975

Promotiecommissie

Promotor:	Prof. dr. F.G. Zitman
Copromotores:	Dr. E.J. Giltay
	Dr. N.J.A. van der Wee
Overige leden:	Prof. dr. J. Gussekloo
	Prof. dr. A.J.W. van der Does
	Prof. dr. A.J.L.M. van Balkom (Vrije Universiteit)

Contents

Chapter 1

General introduction and thesis outline	9
---	---

Chapter 2

Gender differences in clinical characteristics in a naturalistic sample of depressed outpatients: the Leiden Routine Outcome Monitoring study	31
---	----

Chapter 3

Pre-adult versus adult onset major depressive disorder in a naturalistic patient sample: the Leiden Routine Outcome Monitoring study	51
--	----

Chapter 4

Prevalence and correlates of lifetime deliberate self-harm and suicidal ideation in naturalistic outpatients: the Leiden Routine Outcome Monitoring study	71
---	----

Chapter 5

Predicting outcome of depression using the depressive symptom profile: the Leiden Routine Outcome Monitoring Study	93
--	----

Chapter 6

Predicting outcomes of mood, anxiety and somatoform disorders: the Leiden Routine Outcome Monitoring Study	113
--	-----

Chapter 7

Summary, discussion and future perspectives	139
---	-----

Chapter 8

Nederlandse samenvatting	162
List of abbreviations	170
List of publications	172
Curriculum vitae	176
Acknowledgements (dankwoord)	178

Chapter 1

General introduction and thesis outline

M.S. van Noorden

N.J.A. van der Wee

F.G. Zitman

E.J. Giltay

*A modified version of this Chapter has been accepted
for publication as: "Routine Outcome Monitoring in
psychiatric clinical practice: backgrounds and overview"*

Introduction

Routine Outcome Monitoring (ROM) is the systematic measurement of treatment outcomes in routine clinical practice. ROM can be used as a tool for both the patient and the clinician in monitoring treatment progress. With ROM, depending on the choice of measurement instruments, detailed information about psychiatric diagnosis, several domains of symptoms and complaints, and psychosocial functioning can be ascertained, in every phase of treatment. Furthermore, on a group level, anonymised ROM data can be used for conducting epidemiological research, as well as for purposes of benchmarking. ROM is a potentially important source of information regarding the *effectiveness* of treatment in daily –or *real-world*– practice, in addition to the available information about *efficacy* of specific interventions derived from Randomised Controlled Trials (RCTs; Ellwood, 1988; Holloway, 2002; Relman, 1988; Zimmerman et al., 2006). Despite these potential advantages, ROM has not yet been broadly implemented in psychiatry (Carlier et al., 2010; de Beurs et al., 2011; Gilbody et al., 2002; Slade, 2002a).

In the Leiden region in the Netherlands, an extensive ROM infrastructure has been developed and implemented since 2002 by a collaboration of the Regional Mental Health Provider (RMHP) Rivierduinen (RD) and the Department of Psychiatry of the Leiden University Medical Center (LUMC), respectively secondary and tertiary psychiatric specialty care settings (de Beurs, et al., 2011). This thesis focused on real-world patients with Mood, Anxiety and Somatoform (MAS) disorders and real-world outcomes in daily practice, by using ROM-data collected in RD and the LUMC. The remaining part of this introductory chapter provides background information and definitions, as well as the main aims and a thesis outline.

Psychiatric diagnosis

An important precondition for a doctor to adequately treat an ill patient, is a reliable and valid diagnosis. This core criterion is true for all areas in medicine (Goodwin et al., 1996). The study of symptoms and occurrence of diseases, and hence the classification and definition of diagnoses, are within the scope of epidemiology. Ideally, knowledge of underlying pathophysiological disturbances is used for disease classification. Usually, a clinician gathers a medical history, physical examination and often laboratory tests and/or imaging tests to obtain a diagnosis (Fauci et al., 2008). Whenever a reliable and valid diagnosis has been established, a treatment plan can be proposed, and informed consent of the

patient has to be obtained. After initiation, the effect of treatment has to be monitored. In theory, treatment effect can be measured in several domains: disease activity in terms of pathological processes or biological parameters, subjective symptoms as experienced by the patient, symptoms observed by the clinician, (psychosocial) functioning, and health-related quality of life (Fauci, et al., 2008; Smith et al., 1997).

Despite major research efforts in psychiatry during the past decades, knowledge about the pathophysiological mechanisms underlying most psychiatric disorders is still limited. This is in contrast with many somatic disorders, where large breakthroughs in understanding of pathophysiology have been accomplished. This lack of knowledge about pathophysiology of aetiology of psychiatric disorders has implications for both diagnosis and monitoring of treatment effect. Firstly, the value of laboratory tests and other biomarkers in psychiatric diagnostics in the individual patient is merely marginal (Quinones et al., 2009). The psychiatrist uses medical history taking, i.e. the patient's report of internal phenomena and the systematic mental-state examination to ascertain the symptoms and complaints of the patient. Instead of laboratory or imaging tests, rating scales that measure psychopathology can be applied. Secondly, the monitoring of treatment effect is limited to standardised rating of symptoms and psychosocial functioning, because at present, no biological parameters (i.e. biomarkers) can be used as measures of disease activity. However, as mentioned above, monitoring outcomes on a routine and standardised basis with ROM has not yet become standard practice in psychiatry. The various reasons for this lack of implementation will be discussed later in this chapter.

Until 1980, no well-defined, international accepted diagnostic criteria existed in psychiatry (Mayes et al., 2005). The need for reliable and valid diagnoses urged the American Psychiatric Association (APA) to introduce the third edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III) in 1980 (American Psychiatric Association, 1980). This document was based on validity field trials in the United States (US) and consensus of an APA task force, and contrary to the first two editions (DSM-I and DSM-II) it comprised detailed descriptions of symptom clusters and diagnostic criteria of psychiatric disorders. Since the introduction of DSM-III, psychiatric disorders are being classified based on the presence of symptoms, providing syndromal diagnoses. This is exactly what the influential German psychiatrist Emil Kraepelin proposed almost a century earlier (Mayes & Horwitz, 2005). The introduction of the DSM-III has caused a revolution in psychiatry as it dramatically increased the possibilities of conducting epidemiological research with results that were internationally applicable. The current version of the DSM, the DSM-IV (introduced in 1994 with a text revision in 2000; DSM-IV-TR) is the result of

ongoing epidemiological research and consensus (American Psychiatric Association, 1994; American Psychiatric Association, 2000). The publication of the DSM-V is due in 2013 (www.DSM5.org). As examples of common MAS disorders, table 1.1 shows adapted DSM-IV-TR criteria for Major Depressive Episode and Social Phobia.

Table 1.1. Examples of common MAS disorders : adapted DSM-IV-TR criteria of Major Depressive Episode and Social Phobia

Major Depressive Episode	Social Phobia
A. ≥ 5 of the following symptoms present ≥ 2 weeks and represent a change from previous functioning ; at least one of the symptoms is either 1 or 2.	A. An intense fear of ≥ 1 social or performance situations in which the person is exposed to unfamiliar people or to possible scrutiny by others. The individual fears that he or she will act in a way (or show anxiety symptoms) that will be humiliating or embarrassing
1. depressed mood	
2. markedly diminished interest or pleasure	B. Exposure to the feared situation almost invariably provokes anxiety, which may take the form of a situationally predisposed panic attack
3. significant weight loss or weight gain, or decrease or increase in appetite	
4. insomnia or hypersomnia	C. The person recognizes that the fear is excessive or unreasonable
5. psychomotor agitation or retardation	
6. fatigue or loss of energy	D. The situation is avoided or else is endured with intense anxiety or distress
7. feelings of worthlessness or excessive or in appropriate guilt	
8. diminished ability to think or concentrate, or indecisiveness	E. The avoidance, or distress in the feared situation interferes significantly with the person's normal routine, occupational (or academic) functioning, or social activities or relationships, or there is marked distress about having the phobia
9. recurrent thoughts of death or suicide	
B. The symptoms do not meet criteria for a mixed episode	F. The fear or avoidance is not due to the direct physiological effects of a substance or a general medical condition and is not better accounted for by another mental disorder
C. The symptoms cause significant distress or impairment in social, occupational, or other important areas of functioning	
D. The symptoms are not due to the direct physiological effects of a substance (e.g. drug of abuse, medication) or a general medical condition (e.g. hypothyroidism)	G. If a general medical condition or another mental disorder is present, the fear in criterion A is unrelated to it
E. The symptoms are not better accounted for by bereavement.	
	Specify if: generalized if the fears include most social situations .

Abbreviations: MAS: Mood Anxiety and Somatoform; DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision.

Measurement in psychiatry

In order to measure or classify psychiatric disorders, psychiatric symptoms preferably have to be assessed in an objective and reproducible, standardised manner (Zitman, 1990). Since most psychiatric symptoms have a large subjective component (e.g. delusive thoughts, hallucinations, disturbed mood, somatic sensations), objective measurement is a challenge. These symptoms are not easily observed or verified by an examiner (Gilbody et al., 2003). The need for objective measurement of psychiatric symptoms has resulted in the development of psychometrics: the science of psychological assessment.

A psychometric test is an instrument designed to produce a quantitative assessment of some psychological attribute(s). According to the psychometric principles, a psychometric test should be valid, reliable and free of bias (Ishak et al., 2002). Validity indicates that the test assesses the true state of the phenomenon being measured, reliability refers to the extent of reproducibility of the test and bias is a systematic error in the design of the test or study, or in data analysis. For use in ROM, a test should also be sensitive to clinically important change over time (Smith et al., 1997). Measurement in psychiatry can take place on the level of syndromal diagnosis, on the level of symptom severity, on the level of psychosocial functioning and on the level of health related quality of life. Ideally, a ROM test battery consists of measurement instruments that cover all these levels (de Beurs et al., 2011).

DSM Diagnostic measurement instruments

Syndromal classifications are potentially less fundamental than classifications that make use of clearly disturbed biological etiological processes, e.g. the detection of tumor cells in cancer or the occurrence of a pathogen in infectious diseases. Nevertheless, the introduction of the DSM-II and its successive worldwide use has greatly facilitated the development of structured diagnostic measurement instruments, necessary for psychiatric epidemiologic research. Until the 1980's, psychiatric epidemiology was hampered by methodological shortcomings, most importantly because of fuzzy definitions of diagnoses and outcomes (Tohen et al., 2000).

The 1980 DSM-III criteria were used for the development of the Diagnostic Interview Schedule (DIS), for use in the first large US community epidemiologic study on mental health: the Epidemiological Catchment Area (ECA) study, sponsored by the US National Institute of Mental Health (NIMH). This structured interview could be administered by lay interviewers because of its closed-ended questions that did not require clinical judgment (Robins et al., 1981; Tohen et al., 2000). Some years later, the Composite International Diagnostic Interview (CIDI) was developed in collaboration with the World Health Organisation (WHO; Robins et al., 1988). A modified version of the CIDI was used in the next large US community epidemiological study: the National Comorbidity Survey (NCS; Kessler et al., 1994). After the development of these structured diagnostic interviews, reliable prevalence estimates of psychiatric disorders were possible. Because of the extensive format of the CIDI, which limited use in clinical practice, Lecrubier and colleagues developed in a European-US collaboration a short validated structured diagnostic instrument: the Mini International Neuropsychiatric Interview-Plus (MINI-Plus). The MINI-Plus was validated versus the CIDI with satisfactory results (Lecrubier, 1997).

Symptomatic and functional measurement instruments

While diagnostic measurement instruments measure DSM diagnosis in a standardised manner, symptomatic and functional measurement instruments measure symptom severity and health status on a functional level. The latter two categories of instruments can be regarded as monitoring instruments, which may be applied at several time-points during treatment to evaluate progress of treatment and disease. Symptom-based scales may be generic or disorder-specific, and self-report or observer-rated.

The Hamilton Depression Rating Scale (HDRS) and Montgomery-Åsberg Depression Rating Scale (MADRS) are examples of well-known disorder-specific observer-rated rating scales that measure symptom severity in Major Depression (Hamilton, 1960; Montgomery, 1979). Of these two scales, the HDRS has been predominantly used in RCTs, but the MADRS seems superior for outpatient use (Uher et al., 2008). The Brief Symptom Inventory (BSI) is an example of a generic self-report rating scale that measures psychopathological symptom severity in several domains, e.g. anxiety, depression, hostility, and somatic complaints (de Beurs et al., 2006; Derogatis et al., 1983).

Domains of functional monitoring instruments include general health status, quality of life (QoL) and social and role functioning. These instruments are ordinarily regarded as QoL scales, although a distinction between QoL and (psychosocial) functioning can be made. An example of a widely used instrument that measures general health status is the self-report Short Form-36 (SF-36; Ware, 1992).

Psychiatric epidemiology

The large-scale epidemiological community studies facilitated by the sophistication of psychiatric diagnosis and the subsequent development and validation of comprehensive diagnostic measurement instruments have provided valuable data on prevalence and incidence of psychiatric disorders in the general population. Furthermore, these studies have described in great detail clinical characteristics and correlates of most psychiatric disorders. A disadvantage of these epidemiological studies, however, is the fact that the responding subjects do not necessarily reflect treatment seeking populations, even when they meet criteria for psychiatric disorders. This may limit generalisability or external validity of these findings to the daily clinical practice in psychiatric specialty care. Nevertheless, these studies have played a major role in the development of the field of psychiatric epidemiology and the development of current psychiatry. Psychiatric epidemiology can be defined as “the study of the distribution and determinants of mental

disorders in specified populations and of the risk factors associated with their onset and course” (Tohen et al., 2000). In analogy with epidemiology in somatic medicine, psychiatric epidemiology can be subdivided in *community* psychiatric epidemiology and *clinical* psychiatric epidemiology.

In contrast with *community epidemiology*, which aims to describe disease phenomena and to estimate prevalence rates of diseases in the general population, the main goals of *clinical epidemiology* are to investigate the effects of presumed causal risk factors on the onset and course of illness in clinical patients, to evaluate the validity of diagnostic tests and to study predictors of treatment response that might be targeted in subsequent interventions (Kessler, 2007; Tohen et al., 2000). In the last three decades, descriptive community psychiatric epidemiologic research, with the ECA and NCS as examples, has prospered. On the other hand, clinical psychiatric epidemiology has remained under-developed as compared to clinical epidemiology in other fields of medicine (Kessler, 2007). This is mostly because of the fundamental problem of establishing psychiatric diagnoses (assessment of caseness) as compared with somatic diagnoses, because of the limited validity of most psychiatric diagnoses. Another reason for this difference is the fact that the treatment of psychiatric disorders is diverse –despite the availability of evidence-based guidelines– making it more difficult to conduct clinical epidemiological research of naturalistic variation in treatment response (Kessler, 2007). Finally, because many patients with psychiatric disorders do not seek treatment, representative descriptive data of psychiatric disorders in the general population are not necessarily applicable to everyday patients in clinical practice (Burger et al., 2007; Kessler, 2007).

Descriptive community epidemiological studies like the ECA and NCS have yielded valuable insights in prevalence rates and phenomenology of psychiatric disorders in the general population. Since these US community studies in the 1980’s, replications have been conducted (National Comorbidity Survey-replication; NCS-R), as well as community studies in Europe. The first and second Netherlands Mental Health Survey and Incidence Study (NEMESIS 1 and 2) are examples of the latter (Bijl et al., 1998; de Graaf et al., 2011). The prevalence rates of most psychiatric disorders appear to be quite consistent over time and across continents (table 1.2).

Contrary to community epidemiological studies, clinical psychiatric epidemiological studies have the important potential of evaluating interventions in daily clinical practice. Kessler (2007) stated that “In addition to studying the aggregate magnitude of treatment effects, clinical epidemiological studies are needed to study the predictors of individual differences in treatment response”. This type of work would ideally involve investigating baseline (i.e., as from the onset of treatment) predictors of course of

illness in broadly representative clinical samples". Examples of large scale, truly naturalistic studies are scarce nowadays. In order to conduct these studies, a ROM infrastructure could be used, in which outcome data of large naturalistic samples are collected.

Randomised Controlled Trials

The development of psychopathology measurement instruments several decades ago has also dramatically increased the possibilities of evaluating treatment efficacy by means of clinical trials. *Randomised Controlled Trials* (RCTs) have widely been accepted as the gold standard in evaluating treatment efficacy in medicine (Atkins et al., 2004; Kaptchuk, 2001). For a new drug to be approved by the regulatory authorities, superior efficacy compared to placebo in RCTs is needed. Indeed, the design of a typical RCT, in which two or more specific interventions with or without a placebo condition are directly compared in a double-blind way in a sample of patients with a specific disease, aims to maximise internal validity of the trial at hand. In other words, whenever an effect is found, it is most likely being explained by the intervention under study because confounding is largely eliminated through the randomisation process. This high level of internal validity can only be realised if both the disease under study is strictly defined, if the intervention is strictly defined, and if the sample is homogeneous in terms of comorbidity and other clinical characteristics. This means that often a large and strict set of inclusion and exclusion criteria are being used in RCTs. If those strict conditions are met, and the sample is large enough to detect clinically meaningful differences, in theory an RCT with maximised *internal validity* will provide the strongest possible evidence for superiority of a certain intervention (Rorty, 1977).

However, in the real world those perfect 'laboratory circumstances' do not exist. The RCT populations usually are highly selected, and suffer from limited *external validity*. Yet, most evidence-based treatments in medicine are largely based on findings from RCTs. Of course, the findings from RCTs are a major leap forward in terms of *evidence based medicine* as compared to the mere descriptive 'clinical expertise' and case studies from the pre-RCT era.

A major disadvantage of RCTs, however, is the possible lack of external validity or generalisability (Licht et al., 1997; Wells, 1999; Zimmerman et al., 2002). In a recent study of our group we found that only 20-25% of our depressive outpatients would meet general inclusion criteria for RCTs (van der Lem et al., 2010). In other words, real-world patients most likely differ from RCT-patients. Apart from evidence about treatment efficacy, RCTs

have also played a role in our knowledge about characteristics of psychiatric disorders (e.g. symptom profiles, comorbidity patterns). Typically, symptomatology of specific disorders has been analysed using only the baseline measurements in large RCT populations, and reports about these clinical characteristics are being published secondary to the main paper describing the primary outcome of the intervention (see for example Marcus et al., 2005 and Zisook et al., 2007).

Routine Outcome Monitoring

Historical perspective of Routine Outcome Monitoring

The limited generalisability of findings from RCTs and population studies to daily clinical practice and the lack of insight in processes and patient's experiences of treatment has inspired Ellwood for his 1988 Shattuck lecture in which he pleaded for "assessing routinely and frequently the health of patients using appropriate reliable and valid measurement instruments and to build large databases with these data" (Ellwood, 1988). He predicted "a new revolution in health care", and stimulated to systematically assess clinical, financial and health outcomes. Although this idea was well received in editorials (Holloway, 2002; Slade, 2002b), recent reviews have shown only a limited number of published studies of routinely assessed outcomes or ROM in psychiatric specialty care (Carlier et al., 2010). Institutions that have adopted ROM usually used a slim test battery (Burgess et al., 2009; Lambert et al., 2001). Several reasons for this lack of routine implementation of ROM in clinical practice are proposed: ROM is costly and time consuming, and requires a relatively complicated technical infrastructure. More important, no consensus exists about the optimal choice of measurement instruments so that outcomes are not easily comparable across clinics and across studies. Probably, parameters like treatment setting, patient population, and the limited availability of measurement instruments free of copyright may contribute to this lack of consensus. Furthermore, the aim of ROM may vary, as several parties have different interests, e.g. policy makers, insurance companies, patients, clinicians and researchers (Carlier et al., 2010; Norquist, 2002).

Aims and methodological issues of Routine Outcome Monitoring

In theory, ROM can provide both clinician and patient with valuable information about symptoms and treatment outcomes in daily clinical practice, and *effectiveness* of treatments in real-world treatment settings. *Evidence-based* treatments are based on efficacy trials, and may not be effective for every patient in the 'real world'. The main aim

of ROM is improvement of the quality of patient care by measuring progress and giving feedback to the patient. Secondary aims of ROM are understanding mechanisms of disease and treatment, establishing cost effectiveness and benchmarking. For understanding the relationship between patients' health status (outcomes), disease status and treatment (process of care) it is necessary to have access to detailed information about the type of treatment (Smith et al., 1997).

If observer-rated measurement scales are being used, it is important that the interviewer is well trained because clinical interpretation of symptoms or complaints is essential for reliable and reproducible results. To increase objectivity, preferably, measurement instruments are applied by an interviewer who is not directly involved in the treatment of the patient. In addition, inter-rater variability between interviewers should be minimised by recurrent training sessions in which calibration takes place.

Ideally, ROM measurement instruments should be clinically relevant, sensitive to change, minimally burdensome to the patient, to the staff, and to the institution in terms of costs of collection and data analysis (Dickey, 2002). This implies that a balanced selection of well-validated measurement instruments free of copyright is to be preferred. It is evident that different instruments may be applied in different patient groups. In the international literature, no consensus exists about the choice of measurement instruments, about the interval of measurement, and about the groups of patients or treatment settings in which ROM may be applied.

Since the data gathering in ROM is naturalistic and observational, no experimental designs can be used if outcome data are routinely assessed. Hence, instead of causal inferences, only correlations can be established on group level (Kessler, 2007). Another methodological issue when analysing ROM data is the problem of confounding and selection bias, since the treatment that a patient receives will often be determined by a number of factors that are related to outcome, such as disease severity (Gilbody et al., 2002).

The use of patient-based measures of health may itself be useful in improving treatment outcomes, because of the possibility to provide feedback of ROM assessments to both patient and clinician. This may enable clinicians to detect problem areas in treatment that would have been missed without the use of data derived from ROM (Carlier et al., 2010; Greenhalgh et al., 1999), and may increase patient's compliance to treatment protocols (Carlier et al., 2010; Hysong, 2009). A limited amount of studies have demonstrated a positive impact of ROM on monitoring treatment, and on the quality of communication between clinician and patient (Carlier et al., 2010; Knaup et al., 2009). In the meta-analysis by Carlier et al., a favourable outcome of feedback by ROM on mental health was found on the short term only.

ROM in Mood, Anxiety and Somatoform disorders

MAS disorders are highly prevalent disorders with a large disease burden (Wittchen et al., 2011). The frequent chronicity of depression contributes substantially to the global burden of disease. By the year 2020, depression is projected to reach the second ranking place of disability-adjusted life years for all ages in both sexes (Murray et al., 1997). In the second NEMESIS study, a representative survey of 6,646 adults in the Netherlands was conducted between 2007 and 2009 (de Graaf et al., 2011). Lifetime as well as 12-month prevalence rates of mood and anxiety disorders were highly comparable with NEMESIS-1, which dated from 1996 (Bijl et al., 1998), and with NCS/NCS-R data (see table 1.2). Prevalence rates of somatoform disorders have not been ascertained in those studies.

MAS disorders show a considerable amount of overlap in diagnostic criteria and probably also share genetic and environmental factors, so these disorders occur frequently as comorbid disorders. MAS disorders are often regarded as ‘common mental disorders’ as opposed to the less prevalent spectrum of psychotic disorders or ‘severe mental disorders’.

Table 1.2. Lifetime prevalence rates of common mood and anxiety disorders across population studies ^a

	NCS ^b	NCS-R ^c	NEMESIS-1 ^d	NEMESIS-2 ^e
Any mood disorder	19.3 (0.7)	20.8 (0.6)	19.0 (0.5)	20.2 (0.6)
Major Depression	17.1 (0.7)	16.6 (0.5)	15.4 (0.4)	18.7 (0.6)
Dysthymic Disorder	6.4 (0.4)	2.5 (0.2)	6.3 (0.3)	1.3 (0.2)
Any anxiety disorder	24.9 (0.8)	28.8 (0.9)	19.3 (0.5)	19.6 (0.7)
Panic Disorder	3.5 (0.5)	4.7 (0.2)	3.8 (0.2)	3.8 (0.3)
Social Phobia	3.3 (0.7)	12.1 (0.4)	7.8 (0.3)	9.3 (0.5)
Generalised Anxiety Disorder	5.1 (0.3)	5.7 (0.3)	2.3 (0.2)	4.5 (0.3)

Abbreviations: NCS(-R): National Comorbidity Survey(-Replication); NEMESIS: Netherlands Mental Health Survey and Incidence Study; DSM: Diagnostic and Statistical Manual of Mental Disorders; (WMH-)CIDI: (World Mental Health-)Composite International Diagnostic Interview.

^a No prevalence rates were ascertained for somatoform disorders.

^b DSM-III-R disorders measured with the CIDI.

^c DSM-IV disorders measured with the WMH -CIDI.

^d DSM-III-R disorders measured with the CIDI.

^e DSM-IV disorders measured with the WMH -CIDI.

Since ROM is time-consuming and intensive, especially when a large test battery is applied, implementation of ROM in a treatment setting of MAS disorders might be more feasible than in a treatment setting of psychotic disorders. After all, psychotic patients may not be able to complete an extensive set of instruments (Mulder et al., 2010). These considerations have resulted in the implementation of ROM in an outpatient treatment setting of MAS disorders in the Leiden Region (Box 1.1).

Aims and outline of this thesis

Aims of this thesis

This thesis focused on several aspects of ROM in outpatients with MAS disorders. The main aims were to investigate clinical aspects –symptom profiles, comorbidity, general health status and psychosocial functioning– in a large cohort of treatment-seeking real-world patients with MAS disorders, and to compare these aspects with data from RCTs and general population studies. In addition, using prospective data, we aimed to assess whether baseline characteristics measured with ROM could predict real-world –or naturalistic- treatment outcomes. The secondary aim of this thesis was to establish the feasibility of using routinely obtained outcome data for clinical epidemiological research purposes.

General description of the patient population

The implementation of ROM (Box 1.1) in the department of psychiatry of LUMC and in the outpatient departments of RD from 2002 has taken place in several phases. In 2004, most locations had incorporated ROM in routine clinical practice. As from 2004, reliable ROM data has been systematically collected for research purposes. From January 1, 2004 until December 31, 2006, a total of 3,798 patients who had been referred for treatment of a MAS disorder to one of the outpatient clinics in LUMC and RD had a baseline ROM assessment, and were included in the first ROM Baseline cohort of the Leiden Routine Outcome Monitoring Study.

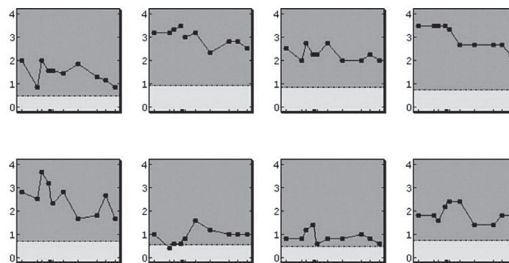
In the first three studies of this thesis, the ROM baseline cohort has been used. The average age of this group was 39.6 (standard deviation [SD] =13.3) years and 63% were women (de Beurs et al., 2011). Diagnostic status was established with the MINI-Plus. It is important to stress that clinical or primary diagnoses have not been taken into account in all ROM analyses. The major advantage of this *modus operandi* is that diagnostic status is entirely based on a validated measurement instrument.

BOX 1.1. ROM in the Leiden University Medical Center & Rivierduinen

In spring 2002, the Regional Mental Health Provider (RMHP) 'Rivierduinen' (an institute serving a region with more than 1 million inhabitants) and the Department of Psychiatry of the Leiden University Medical Center (LUMC) started collaboration for routine assessment of the DSM-IV diagnosis as well as the symptom severity, well-being and health status at time of the first interview of outpatients referred to the RMPH Rivierduinen.

At the start, ROM was restricted to patients referred for treatment of mood, anxiety, and somatoform (MAS) disorders. These patients form a relatively homogenous group with substantial mutual comorbidity (Kessler et al., 1996) and they mainly receive outpatient care. To be eligible, patients had to have sufficient mastery of the Dutch language and had to be able to complete self-report instruments. Patients who are considered (by their clinician) to be too ill to complete questionnaires or refuse to be assessed are excluded from ROM assessment.

All patients are assessed by an independent psychiatric research nurse at the start, and during follow up at intervals of three to four months, at the beginning of a new treatment step and at the end of the treatment.



A disadvantage may be that in the case of multiple diagnoses, one diagnosis could be more prominent, and treatment may be based on this clinically more relevant diagnosis. For example, a patient with a Major Depressive Disorder (MDD) and a social phobia, may either be treated according to the social phobia treatment or the major depression guidelines. Until now, no detailed information about treatment is ascertained in the Leiden ROM. Hence, this particular patient may be analysed both as a MDD case and a social phobia case (or as patient with comorbidity), irrespective of the primary diagnosis and of treatment.

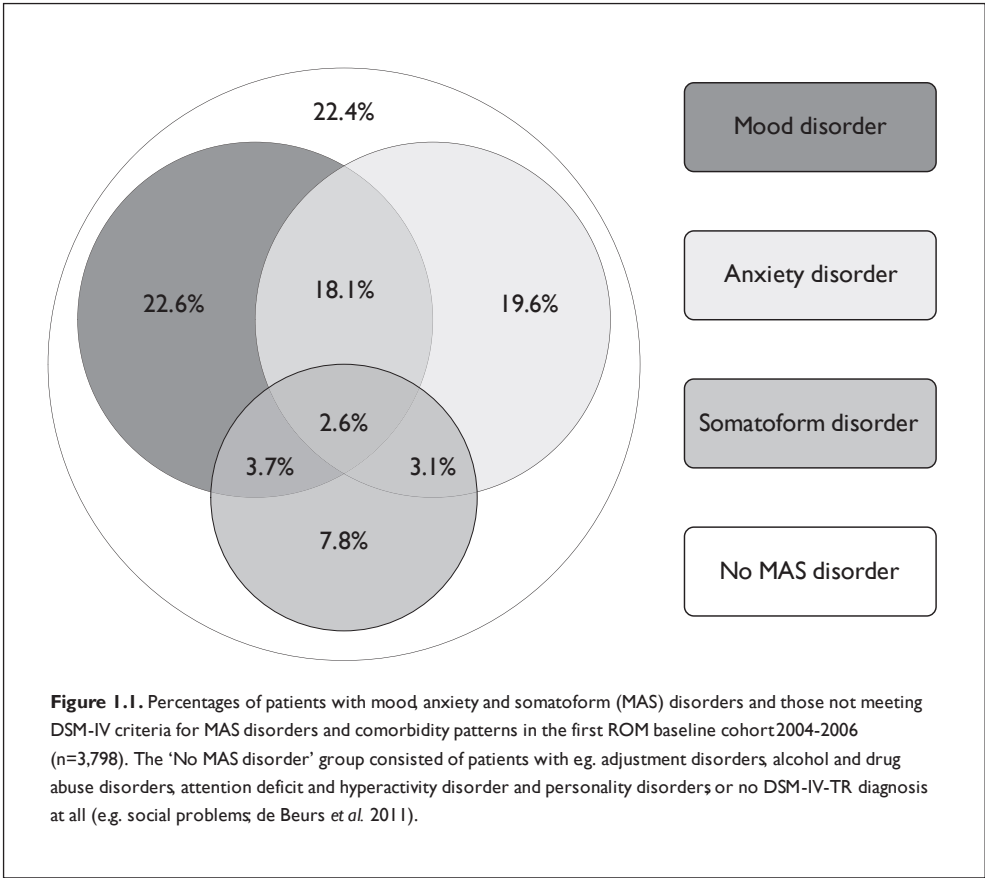
BOX 1.1. Continued

During the first session, a standardised diagnostic interview is administered and observer- and self-reported ratings are determined. At baseline the Axis-I diagnosis according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) is established using the Mini-International Neuropsychiatric Interview-plus (MINI-plus, Sheehan et al., 1998). The interviews are performed by psychiatric research nurses who have been extensively trained and supervised. The Dimensional Assessment of Personality Pathology (DAPP-SF) is administered to assess maladaptive personality traits (Livesley et al., 2006; van Kampen et al., 2008). Until now, in ROM no detailed treatment information is available.

Subsequently, a number of symptom severity rating scales are administered at baseline, which are also completed at each re-assessment to allow for the evaluation of treatment outcome. Together, these instruments cover change in three areas of functioning: symptom reduction, increased wellbeing, and improvement in general life functioning (Sperry et al., 1996). They are commonly used in treatment-outcome research and have good psychometric properties as evidenced by national and international publications (an overview of instruments used is available at <http://www.lumc.nl/psychiatry/ROM-instruments>). Outcome is assessed by patients' self-report and by an independent assessor (observer-rated), and includes both generic and disorder-specific measures. Clinicians receive a report on the results of the baseline assessments as well as follow-up reporting on treatment outcome in the above mentioned domains. Results of the assessments are provided in detail by the research nurses as well as in a summarised form. The summaries facilitate clinicians to discuss the results with their patients and use them as a tool to evaluate the treatment. Results are also used, in an anonymous form, for scientific purposes.

Since ROM-data are primarily being used by clinicians and patients to monitor treatment progress, no specific informed consent is needed. The use of anonymised data for research purposes has been approved by the Medical Ethical Committee of the LUMC.

According to the MINI-Plus, of the 3,798 patients, 1,788 patients (47.0%) met criteria for one or more mood disorders, 1,653 patients (43.5%) had one or more anxiety disorders, 653 patients (17.2%) had one or more somatoform disorders, and 851 patients (22.4%) had no MAS disorder (figure 1.1). These latter patients had other diagnoses, e.g. adjustment disorders, attention deficit hyperactivity disorder, substance abuse disorders, or did not meet criteria for a Axis I DSM-IV diagnosis (de Beurs et al., 2011).



Thesis outline

In *Chapter 2*, we focused on gender differences in point prevalence rate, symptom profile and comorbidity patterns in 1,131 outpatients with MDD. The aim of this exploratory study was to investigate whether gender differences in a real-world sample of treatment-seeking MDD patients existed and whether these differences would deviate from differences found in clinical trial or population-based samples.

Chapter 3 describes a study on differences in clinical characteristics including demographic correlates, comorbidity, symptomatology and general health status in patients with MDD onset before adulthood versus patients with MDD onset during adulthood. Again, results were compared with findings from non-naturalistic studies.

In *Chapter 4*, we shifted the focus to lifetime deliberate self-harm and suicidal ideation (DSHI) in the total group of MAS patients. It is well-known that self-harm is a common problem in psychiatric specialty care, but no large-scale studies have described

correlates of DSHI in a real-world sample of outpatients with common mental disorders. The aim of this study was to investigate prevalence and correlates of DSHI in a large real-world outpatient mental health population.

In *Chapter 5*, we used longitudinal ROM data to identify predictors of outcome in MDD. Specifically, we aimed to investigate whether individual depressive symptoms measured with the Beck Depression Inventory-Revised (BDI-II) predict response or remission after follow-up of up to 24 months.

Chapter 6 describes a study on predictors of outcome in MAS outpatients using longitudinal ROM data. Since MAS disorders often have overlapping symptoms and mutual comorbidity is substantial, we investigated cross-diagnostic predictors of outcome measured on common generic rating scales: the BSI and the Clinical Global Impression (CGI) scale.

Finally, in *Chapter 7*, we summarised the main results of the studies. Furthermore, we discussed these results and we provided recommendations for further improvement of ROM and future studies.

References

- American Psychiatric Association, 1980. Diagnostic and statistical Manual of Mental Disorders. Washington, DC.
- American Psychiatric Association, 1994. Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised. Washington, DC.
- American Psychiatric Association, 2000. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. Washington, DC.
- Atkins, D., Best, D., Briss, P.A., Eccles, M., Falck-Ytter, Y., Flottorp, S., Guyatt, G.H., Harbour, R.T., Haugh, M.C., Henry, D., Hill, S., Jaeschke, R., Leng, G., Liberati, A., Magrini, N., Mason, J., Middleton, P., Mrukowicz, J., O'Connell, D., Oxman, A.D., Phillips, B., Schunemann, H.J., Edejer, T.T., Varonen, H., Vist, G.E., Williams, J.W., Jr., Zaza, S., 2004. Grading quality of evidence and strength of recommendations. *British Medical Journal* 328, 1490.
- de Beurs, E., den Hollander-Gijsman, M.E., van Rood, Y.R., van der Wee, N.J., Giltay, E.J., van Noorden, M.S., van der Lem, R., van Fenema, E.M., 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 and Psychotherapy* 18, 1-12.
- de Beurs, E., Zitman, F.G., 2006. The Brief Symptom Inventory (BSI): Reliability and validity of a practical alternative to SCL-90. *Maandblad Geestelijke Volksgezondheid* 61, 120-141.
- Bijl, R.V., Ravelli, A., van Zessen, G., 1998. Prevalence of psychiatric disorder in the general population: results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology* 33, 587-595.
- Burger, H., Neeleman, J., 2007. A glossary on psychiatric epidemiology. *Journal of Epidemiology and Community Health* 61, 185-189.
- Burgess, P.M., Pirkis, J.E., Slade, T.N., Johnston, A.K., Meadows, G.N., Gunn, J.M., 2009. Service use for mental health problems: findings from the 2007 National Survey of Mental Health and Wellbeing. *Australian and New Zealand Journal of Psychiatry* 43, 615-623.
- Carlier, I.V., Meuldijs, D., van Vliet, I.M., van Fenema, E.M., van der Wee, N.J., Zitman, F.G., 2010. Routine outcome monitoring and feedback on physical or mental health status: evidence and theory. *Journal of Evaluation in Clinical Practice* (in press).
- Derogatis, L.R., Melisaratos, N., 1983. The Brief Symptom Inventory: an introductory report. *Psychological Medicine* 13, 595-605.
- Dickey, B., 2002. Outcome measurement from research to clinical practice. In: Ishak, W.W., Burt, T., Sederer, L.I. (eds.), *Outcome measurement in psychiatry*. American Psychiatric Publishing, Washington, DC.
- Ellwood, P.M., 1988. Shattuck lecture: Outcomes management. A technology of patient experience. *New England Journal of Medicine* 318, 1549-1556.
- Fauci, A.S., Braunwald, E., Kasper, D.L., Hauser, S.L., Longo, D.L., Jameson, D.L., Loscalzo, J., 2008. *The Practice of Medicine*. In: Fauci, A.S., Braunwald, E., Kasper, D.L., Hauser, S.L., Longo, D.L., Jameson, D.L., Loscalzo, J. (eds.), McGraw-Hill, New York.
- Gilbody, S., House, A.O., and Sheldon, T.A. Outcomes measurement in psychiatry: a review of outcomes management in psychiatric research and practice. Centre for Reviews and Dissemination report 24 . 2003. University of New York, New York.

- Gilbody, S.M., House, A.O., Sheldon, T.A., 2002. Outcomes research in mental health. Systematic review. *The British Journal of Psychiatry* 181, 8-16.
- Goodwin, D., Guze, S., 1996. *Psychiatric diagnosis*. Oxford University Press, New York.
- de Graaf, R., Ten Have, M., van Gool, C., van Dijk, S., 2011. Prevalence of mental disorders and trends from 1996 to 2009. Results from the Netherlands Mental Health Survey and Incidence Study-2. *Social Psychiatry and Psychiatric Epidemiology* (in press).
- Greenhalgh, J., Meadows, K., 1999. The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review. *Journal of Evaluation in Clinical Practice* 5, 401-416.
- Hamilton, M., 1960. A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry* 23, 56-62.
- Holloway, F., 2002. Outcome measurement in mental health--welcome to the revolution. *The British Journal of Psychiatry* 181, 1-2.
- Hysong, S.J., 2009. Meta-analysis: audit and feedback features impact effectiveness on care quality. *Medical Care* 47, 356-363.
- Ishak, W.W., Burt, T., Sederer, L.I., 2002. *Outcome Measurement in Psychiatry: a critical review*. American Psychiatric Association, Washington, DC.
- 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.
- Kaptchuk, T.J., 2001. The double-blind, randomised, placebo-controlled trial: gold standard or golden calf? *Journal of Clinical Epidemiology* 54, 541-549.
- Kessler, R.C., 2007. Psychiatric epidemiology: challenges and opportunities. *International Review of Psychiatry* 19, 509-521.
- Kessler, R.C., McGonagle, K.A., Zhao, S.Y., Nelson, C.B., Hughes, M., Eshleman, S., Wittchen, H.U., Kendler, K.S., 1994. Lifetime and 12-Month Prevalence of Dsm-III-R Psychiatric-Disorders in the United-States - Results from the National-Comorbidity-Survey. *Archives of General Psychiatry* 51, 8-19.
- Kessler, R.C., Nelson, C.B., McGonagle, K.A., Liu, J., Swartz, M., Blazer, D.G., 1996. Comorbidity of DSM-III-R major depressive disorder in the general population: Results from the US National Comorbidity Survey. *The British Journal of Psychiatry* 168, 17-30.
- Knap, C., Koesters, M., Schoefer, D., Becker, T., Puschner, B., 2009. Effect of feedback of treatment outcome in specialist mental healthcare: meta-analysis. *The British Journal of Psychiatry* 195, 15-22.
- Lambert, M.J., Hansen, N.B., Finch, A.E., 2001. Patient-focused research: using patient outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology* 69, 159-172.
- Lecrubier, Y., 1997. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *European Psychiatry* 12, 224-231.
- van der Lem, R., van der Wee, N.J., van Veen, T., Zitman, F.G., 2010. The generalisability of antidepressant efficacy trials to routine psychiatric out-patient practice. *Psychological Medicine* 41, 1353-1363.
- Licht, R.W., Gouliaev, G., Vestergaard, P., Frydenberg, M., 1997. Generalisability of results from randomised drug trials. A trial on antimanic treatment. *The British Journal of Psychiatry* 170, 264-267.

- Livesley, W.J., Jackson, D.N., 2006. Manual for the Dimensional Assessment of Personality Problems -basic questionnaire. Sigma, Port Huron, Michigan.
- Marcus, S.M., Young, E.A., Kerber, K.B., Kornstein, S., Farabaugh, A.H., Mitchell, J., Wisniewski, S.R., Balasubramani, G.K., Trivedi, M.H., Rush, A.J., 2005. Gender differences in depression: findings from the STAR*D study. *Journal of Affective Disorders* 87, 141-150.
- Mayes, R., Horwitz, A.V., 2005. DSM-III and the revolution in the classification of mental illness. *Journal of the History of the Behavioural Sciences* 41, 249-267.
- Montgomery, S.A., 1979. A new depression scale designed to be sensitive to change. *The British journal of psychiatry* 134, 382-389.
- Mulder, C.L., Van der Gaag, M., Bruggeman, R., Cahn, W., Delespaul, P.A., Dries, P., Faber, G., de Haan, L., van der Heijden, F.M., Kempen, R.W., Mogendorff, E.S., Slooff, C.J., Sytema, S., Wiersma, D., Wunderink, L., van Os, J., 2010. [Routine Outcome Monitoring for patients with severe mental illness: a consensus document]. *Tijdschrift voor Psychiatrie* 52, 169-179.
- Murray, C.J., Lopez, A.D., 1997. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet* 349, 1498-1504.
- Norquist, G.S., 2002. Role of outcome measurement in psychiatry. In: Ishak, W.W., Burt, T., Sederer, L.I. (eds.), *Outcome measurement in psychiatry*. American Psychiatric Publishing, Washington, DC.
- Quinones, M.P., Kaddurah-Daouk, R., 2009. Metabolomics tools for identifying biomarkers for neuropsychiatric diseases. *Neurobiology of Disease* 35, 165-176.
- Relman, A.S., 1988. Assessment and accountability: the third revolution in medical care. *New England Journal of Medicine* 319, 1220-1222.
- Robins, L.N., Helzer, J.E., Croughan, J., Ratcliff, K.S., 1981. National Institute of Mental Health Diagnostic Interview Schedule. Its history, characteristics, and validity. *Archives of General Psychiatry* 38, 381-389.
- Robins, L.N., Wing, J., Wittchen, H.U., Helzer, J.E., Babor, T.F., Burke, J., Farmer, A., Jablenski, A., Pickens, R., Regier, D.A., 1988. The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry* 45, 1069-1077.
- Rorty, R., 1977. *Philosophy and the mirror of nature*. Princeton University Press, Princeton.
- Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry* 59 Suppl 20, 22-33.
- Slade, M., 2002a. What outcomes to measure in routine mental health services, and how to assess them: a systematic review. *Australian and New Zealand Journal of Psychiatry* 36, 743-753.
- Slade, M., 2002b. Routine outcome assessment in mental health services. *Psychological Medicine* 32, 1339-1343.
- Smith, G.R., Jr., Manderscheid, R.W., Flynn, L.M., Steinwachs, D.M., 1997. Principles for assessment of patient outcomes in mental health care. *Psychiatric Services* 48, 1033-1036.
- Sperry, L., Brill, P.L., Howard, K.I., Grissom, G.R., 1996. *Treatment outcomes in psychotherapy and psychiatric interventions*. Brunner/Mazel, Philadelphia, PA.
- Tohen, M., Bromet, E., Murphy, J.M., Tsuang, M.T., 2000. Psychiatric epidemiology. *Harvard Review of Psychiatry* 8, 111-125.

- Uher, R., Farmer, A., Maier, W., Rietschel, M., Hauser, J., Marusic, A., Mors, O., Elkin, A., Williamson, R.J., Schmael, C., Henigsberg, N., Perez, J., Mendlewicz, J., Janzing, J.G., Zobel, A., Skibinska, M., Kozel, D., Stamp, A.S., Bajs, M., Placentino, A., Barreto, M., McGuffin, P., Aitchison, K.J., 2008. Measuring depression: comparison and integration of three scales in the GENDEP study. *Psychological Medicine* 38, 289-300.
- Ware Jr, J.E., 1992. The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection. *Medical Care* 30, 473.
- Wells, K.B., 1999. Treatment Research at the Crossroads: The Scientific Interface of Clinical Trials and Effectiveness Research. *The American Journal of Psychiatry* 156, 5.
- Wittchen, H.U., Jacobi, F., Rehm, J., Gustavsson, A., Svensson, M., Jonsson, B., Olesen, J., Allgulander, C., Alonso, J., Faravelli, C., Fratiglioni, L., Jennum, P., Lieb, R., Maercker, A., van Os, J., Preisig, M., Salvador-Carulla, L., Simon, R., Steinhausen, H.C., 2011. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European Neuropsychopharmacology* 21, 655-679.
- Zimmerman, M., Mattia, J.I., Posternak, M.A., 2002. Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? *The American Journal of Psychiatry* 159, 469-473.
- Zimmerman, M., Ruggero, C.J., Chelminski, I., Young, D., Posternak, M.A., Friedman, M., Boerescu, D., Attiullah, N., 2006. Developing brief scales for use in clinical practice: the reliability and validity of single-item self-report measures of depression symptom severity, psychosocial impairment due to depression, and quality of life. *Journal of Clinical Psychiatry* 67, 1536-1541.
- Zisook, S., Lesser, I., Stewart, J.W., Wisniewski, S.R., Balasubramani, G.K., Fava, M., Gilmer, W.S., Dresselhaus, T.R., Thase, M.E., Nierenberg, A.A., Trivedi, M.H., Rush, A.J., 2007. Effect of age at onset on the course of major depressive disorder. *The American Journal of Psychiatry* 164, 1539-1546.
- Zitman, F.G., 1990. Standaardisering van psychiatrische diagnostiek. In: Abraham, R.E., Giel, R., Rooijmans, H.G.M., Zitman, F.G. (eds.), *Diagnostiek in de psychiatrie, mogelijkheden en grenzen*. Boerhaave Commissie, Leiden.

Chapter 2

Gender differences in clinical characteristics
in a naturalistic sample of depressive outpatients:
the Leiden Routine Outcome Monitoring Study

Martijn van Noorden
Erik Giltay
Margien den Hollander-Gijsman
Nic van der Wee
Tineke van Veen
Frans Zitman

Journal of Affective Disorders (2010), 125, 116-23

Abstract

Background: No previous large scale studies have assessed gender differences in naturalistic samples of major depressive disorder (MDD) outpatients. We therefore determined gender differences in comorbidity, symptom patterns and subjective health status in these outpatients in a mental healthcare setting.

Methods: Of 3,798 consecutive adult patients (age range: 18-65), 1,131 (65.1% women) fulfilled DSM-IV criteria of current MDD on the Mini International Neuropsychiatric Interview (MINI-Plus). Patients were routinely assessed with Routine Outcome Monitoring (ROM), including the Montgomery-Åsberg Depression Rating Scale (MADRS), Beck Depression Inventory (BDI-II), Brief Symptom Inventory (BSI) and Short Form-36 (SF-36).

Results: No gender differences were found in disease severity using the clinician-rated MADRS. However, women showed a significant higher depression severity measured with the self-report BDI-II. Also, psychopathological symptoms self-reported with the BSI were higher, and reported health status on the SF-36 was lower in women. In men with MDD, social phobia, attention deficit hyperactivity disorder, and alcohol and drug misconduct were more common co-morbid disorders, while in women with MDD posttraumatic stress disorder and bulimia nervosa were more common, as well as atypical features of depression. Limitations: The use of retrospective reports of lifetime psychopathology might have led to recall bias. 20% of subjects were excluded from ROM due to language problems or logistical reasons.

Discussion: Although women self-reported higher depression severity, more severe general psychopathological symptoms and lower health status, no differences in disease severity were found on interviewer ratings. These findings could have implications for clinical decision making and treatment.

Introduction

Although differences in course and symptom patterns between male and female depressive patients are not reflected in classification systems, there is evidence that important gender differences exist that point to underlying differences in the pathophysiology of depression (Smith et al., 2008). Gender differences in the prevalence of major depressive disorder (MDD) have extensively been reported based on both population-based (Angst et al., 2002; Kessler et al., 2005; Kuehner, 2003) and clinical trial samples (Fava et al., 1996; Kornstein et al., 2000; Marcus et al., 2005; Marcus et al., 2008). Consistently, lifetime prevalence ratios are 2:1 for women as compared to men, whereas a ratio of 1.7:1 for point prevalence rates is reported (Angst et al., 2002; Kessler et al., 2003; Kuehner, 2003; Nolen-Hoeksema, 1987). The cause of the female preponderance of MDD remains to be elucidated (Kendler, 2001; Kuehner, 2003; Middeldorp et al., 2006; Piccinelli & Wilkinson, 2000). In addition to gender differences in prevalence rates, differences in the course, comorbidity patterns, clinical characteristics, and general health status between men and women with MDD have been studied. Studies on these gender differences in clinical manifestations of MDD have been conducted in population-based samples and in groups of MDD-patients who were recruited for randomised controlled trials (RCTs), but scarcely in naturalistic samples in psychiatric specialty care. One study on gender differences in severity and symptomatology of depression in a representative sample of depressive patients in general practice showed no clinically relevant gender differences (Hildebrandt, 2003). A recent study in a genetic-epidemiological cohort of 598 MDD-patients, however, showed differences in the course and symptom profile of male and female depression with more atypical depression, self-reproach and diminished libido in women (Smith et al., 2008).

Large population samples of six European countries were analysed in the European DEPRES study (Angst et al., 2002). In this study of more than 78,000 randomly selected subjects (38,434 males and 40,024 females), the sex-specific prevalence rates of all nine DSM-IV symptoms of depression were scored in a diagnostic interview. 2,921 men (7.6%) and 5,968 women (14.9%) were diagnosed with MDD. All symptoms were more prevalent in women than in men. In addition, women were more likely to seek help for their complaints. Overall, men reported fewer symptoms than women during a clinical interview, which made the authors conclude that the preponderance of women with a major depressive episode could be a result of the diagnostic (DSM-IV) threshold of 5 or more symptoms used in this study. Indeed, when the criteria for 'minor depression' were applied, namely two symptoms instead of four symptoms in addition to one core

symptom, the gender ratio was much lower (female to male ratio 1.2: 1). In the literature, this phenomenon is regarded as 'threshold artefact' (Piccinelli & Wilkinson, 2000). In this study, however, it was unclear whether men with MDD would also report less severe depressive symptoms than women with MDD.

Studies in clinical trial samples were relatively small (Fava et al., 1996; Frank et al., 1988; Grigoriadis et al., 2007), with the exception of the STAR*D study samples (Rush et al., 2004). In this US outpatient treatment study, gender differences in baseline sociodemographic and clinical characteristics were studied in two cohorts of 1,500 and 2,541 MDD patients respectively, after informed consent was provided (Marcus et al., 2005; Marcus et al., 2008). Contrary to other treatment studies, the inclusion criteria were not too strict and equipoise randomisation was used in order to reflect a real-life patient population. Depressed women had a greater overall symptom severity measured on observational and self-report scales, and were more likely to have a comorbid anxiety disorder, bulimia nervosa or somatoform disorder, similar to previous findings (Kuehner, 2003; Middeldorp et al., 2006).

To our knowledge, no data on gender differences in large naturalistic samples of depressed outpatients in routine psychiatric care have been published. Therefore we studied gender differences in point prevalence rate, symptom profile and comorbidity pattern of MDD in patients enrolled in the Leiden Routine Outcome Monitoring (ROM) Study database. The extensive phenotyping provided a unique opportunity to analyse these gender differences in a large naturalistic sample of patients with MDD in secondary and tertiary routine psychiatric care.

Methods

Routine Outcome Monitoring

We used a sample of 3,798 adults (age range 18-65) who were referred for treatment of a mood, anxiety, or somatoform (MAS) disorder to the Dutch Regional Mental Health Provider (RMHP) Rivierduinen (RD) or the psychiatric outpatient department of the Leiden University Medical Center (LUMC) in the Western part of the Netherlands between January 2004 and December 2006. At baseline, subjects were assessed as part of the usual Routine Outcome Monitoring (ROM) procedure. In ROM, all patients referred to RD or LUMC for treatment of a mood, anxiety or somatoform disorder are routinely assessed with an extensive psychometric battery at baseline and prospectively during treatment. Only patients with insufficient mastery of the Dutch language and patients who are unable to

complete computerised and written questionnaires are ineligible for ROM. On average, 80 percent of the referred patients with a tentative MAS disorder were assessed with ROM in the study period. ROM data are primarily used for diagnosis and to inform clinicians and patients about treatment progress. Informed consent is not required.

For our study, we used only the baseline ROM assessments. These baseline ROM assessments comprised a standardised diagnostic interview (Dutch version of the Mini International Neuropsychiatric Interview-Plus, version 5.00-R; MINI-Plus) focussing on DSM-IV-TR diagnosis (Sheehan et al., 1998; van Vliet & de Beurs, 2007), collection of sociodemographic and socioeconomic data, observer-rated scales and self-report questionnaires, and general measures of health and quality of life in order to assess psychopathology dimensionally as well as categorically. Both generic and disorder-specific scales were used. The observational scales were completed in a face-to-face interview, whereas the self-report questionnaires were filled out by the patient using a touch-screen computer. The assessments were performed by trained research nurses in the outpatient clinics. The use of anonymised data for research purposes has been approved by the Medical Ethical Committee of the Leiden University Medical Hospital.

We included all patients with a current DSM-IV-TR MDD according to the MINI-Plus in the present analyses. Patients with incomplete data, a bipolar disorder or lifetime psychotic illness were excluded, but a diagnosis of psychotic depression was allowed. The MINI-Plus was shown to have a sensitivity of 0.94 and a specificity of 0.79 in identifying MDD. Furthermore, the MINI-Plus has been validated with Composite International Diagnostic Interview (CIDI) diagnoses (World Health Organisation, 1990), and a good reliability has been shown (Lecrubier, 1997). 1637 patients with a current MDD according to the MINI-Plus were identified (43.1%). Complete data on all variables of interest were obtained from 1131 of these patients (69.1%) aged 18-65 years, of whom 53 (4.7%) were tertiary care patients. Patients who were missing one or more variables did not differ significantly from the patients with complete data coverage of the outcome variables (data not shown).

Variables

Demographic variables were obtained using a self-report questionnaire that assessed ethnic background, education, marital status, housing situation and employment status. A Dutch ethnic background was assumed when the patient and both parents were born in the Netherlands. Subjective depression severity and individual symptoms were assessed with the Beck Depression Inventory-revised (BDI-II), a 21-item self-report instrument with good psychometric properties (Beck et al., 1988, Beck et al., 1996). Individual symptoms were defined as present if the patient scored 2 or 3 on a 4-point likert-scale ranging from

0 to 3. The two items that assess eating and sleeping activity were recoded into four variables: insomnia, hypersomnia, anorexia and hyperphagia.

The abbreviated Comprehensive Psychopathological Rating Scale consists of the Montgomery-Åsberg Depression Rating Scale (MADRS; Montgomery, 1979), the Brief Anxiety Scale (BAS; Tyrer et al., 1984) and a scale that assesses psychomotor inhibition (REM; Goekoop, 1992). In our study, objective depression severity was assessed with the MADRS. The MADRS has a good internal consistency and reliability (Montgomery, 1979). In addition to the continuous measure, we distinguished a MADRS score of <20 for mild depression, 20-35 for moderate depression, and ≥ 35 for severe depression (Muller, 2003; Snaith, 1986).

The Brief Symptom Inventory (BSI) is a short version of the Symptom Checklist (SCL-90), a self-report instrument that assesses psychopathological symptoms in several domains, e.g. somatic symptoms, depressive symptoms, anxiety symptoms and hostility (de Beurs & Zitman, 2006; Derogatis & Melisaratos, 1983). The BSI has shown good internal consistency, reliability and validity (Derogatis, 1977).

If a patient had a MDD according to the MINI-Plus, the patients age at the time of onset of the first MDD episode was assessed by asking: "how old were you when you first experienced these symptoms of depression?" Subsequently, the number of MDD episodes was asked for. Both prior suicide attempts and current suicidal thoughts (during the past month) were assessed with the MINI-Plus.

Melancholic features of depression were assessed with the MINI-Plus according to the DSM-IV criteria. Since the MINI-Plus does not systematically assess the DSM-IV atypical features of depression, we considered atypical depression present when patients had two or more symptoms of polyphagia, fatigue and hypersomnia according to the BDI-II items, as previously used by Angst et al. (2006)

Generic health status was assessed with the Dutch version of the Short Form-36 Health Survey (SF-36), a 36 item self-report questionnaire that measures health status in eight domains (Aronson, 1998; Ware, 1992).

Statistical analyses

Chi squared tests were used to compare categorical variables and t-tests for independent samples to compare continuous measures between men and women. Bonferroni-correction was performed in post-hoc tests. In case of a skewed distribution of the data, square-root and natural logarithm transformations were applied to obtain a normal distribution before the use in statistical testing. Consecutively, data were back-transformed, and geometric means with 5% and 95% percentiles are given in the tables. Adjustments were made for age and ethnic background in logistic regression models

(for discrete variables). Alternatively, analysis of covariance (ANCOVA, for continuous variables) was used. Comorbidity and individual BDI-II symptoms were adjusted for the MADRS score which reflect depressive symptom severity. Significance level was set at $p < 0.05$. Data were analysed using SPSS version 16.0 statistical software.

Results

Table 2.1. Sociodemographic characteristics according to gender in 1,131 outpatients with current MDD

	Men (n=395; 34.9%)	Women (n=736; 65.1%)	p-value
Age (geometric mean; P5-P95)	40.0 (21-59)	36.2 (20-58)	<0.001
Ethnic background (n, %)			0.54
Dutch	310 (78.5)	589 (80.0)	
Other ethnicity	85 (21.5)	147 (20.0)	
Marital status (n, %)			0.90
Married/living with partner	190 (48.1)	350 (47.6)	
Unmarried	205 (51.9)	386 (52.4)	
Housing situation (n, %)			<0.001
Living alone	133 (33.7)*	172 (23.4)*	
Living with partner	191 (48.4)	355 (48.2)	
Living with family	71 (18.0)*	209 (28.4)*	
Educational status (n, %)			0.61
Lower education	53 (13.4)	85 (11.5)	
High school (lower)	126 (31.9)	261 (35.5)	
High school (higher)	151 (38.2)	274 (37.2)	
College/university	65 (16.5)	116 (15.8)	
Employment status (n, %)			<0.001
Employed - part time	27 (6.8)*	172 (23.4)*	
Employed - full time	119 (30.1)*	90 (12.2)*	
Unemployed/retired	84 (21.3)*	232 (31.5)*	
Work-related disability	165 (41.8)*	242 (32.9)*	

MDD denotes Major Depressive Disorder .

* indicates that percentage is significantly different in post-hoc comparisons after Bonferroni correction.

Sample and demographic data

Table 2.1 shows the sociodemographic characteristics of the subjects. The sample consisted of 395 men (34.9%) and 736 women (65.1%), a male to female ratio of 1 to 1.86. The mean age at assessment was 3.8 years higher in men than in women ($p < 0.001$). Ethnic

background did not differ significantly between men and women. 21.5% of men and 20.0% of women were of non-Dutch origin. More men reported to be living alone than women. Women worked part-time significantly more frequently than men, and men were more likely to have full-time employment.

Rating scale scores and clinical characteristics (Table 2.2)

Women reported more severe depressive symptom scores on the BDI-II than men (BDI-II total scores of 32.6 vs. 29.3., respectively, $p < 0.001$ adjusted for age and ethnic background). Also, the number of BDI-II symptoms (presence defined as a score of 2 or 3 on the 0-3 point Likert-scale) was higher in women than in men (10.36 out of 21 items vs. 8.89 out of 21 items, respectively, $p < 0.001$, data not shown). However, depressive symptom severity assessed by the research nurse on the observational MADRS did not differ between men and women. Mild (MADRS < 20), moderate (MADRS 20-35) and severe depressive symptoms (MADRS ≥ 35) were distributed equally over the sexes. No significant differences were found in severity of anxiety symptoms measured with the observational BAS, but observed psychomotor inhibition was significantly more pronounced in men than in women.

Table 2.2. Rating scale scores and clinical characteristics according to gender in 1,131 outpatients with current MDD

	Men (n=395; 34.9%)	Women (n=736; 65.1%)	p-value ^a	Adjusted p-value ^b
BDI-II score (mean $\pm$ SD)	29.3 $\pm$ 10.5	32.6 $\pm$ 10.5	<0.001	<0.001
MADRS score (mean $\pm$ SD)	25.1 $\pm$ 8.2	25.3 $\pm$ 7.4	0.60	0.22
MADRS (n, %)			0.41	0.37
<20	85 (21.5)	143 (19.5)		
20-35	259 (65.4)	509 (69.3)		
≥ 35	52 (13.1)	83 (11.3)		
BAS score (mean $\pm$ SD)	16.9 $\pm$ 6.8	17.5 $\pm$ 6.3	0.14	0.09
REM score (geometric mean, P5-P95)	5.2 (1-12)	4.6 (1-12)	0.005	0.03
BSI total score (geometric mean, P5-P95)	1.4 (0.5-2.8)	1.5 (0.5-2.9)	0.01	0.03
Age of onset of first MDD episode (geometric mean, P5-P95)	28.9 (15-57)	25.4 (14-53)	<0.001	0.06
Number of MDD episodes (n, %)			0.20	0.12
1	159 (40.3)	273 (37.1)		
2-4	115 (29.1)	253 (34.4)		
>4	121 (30.6)	210 (28.5)		
Prior suicide attempt (n, %)	72 (18.2)	174 (23.6)	0.04	0.04
Current suicidal thoughts (n, %)	150 (38.0)	269 (36.5)	0.65	0.42

Abbreviations: MDD: Major Depressive disorder ; BDI-II: Beck Depression Inventory – revised; MADRS: Montgomery-Åsberg Depression Rating Scale; BAS: Brief Anxiety Scale; REM: Comprehensive Psychiatric Rating Subscale (psychomotor inhibition); BSI: Brief Symptom Inventory.

^a p-value by ANOVA or chi-squared test, when appropriate;

^b adjusted for age and ethnic background.

Women displayed a higher score than men on the BSI self-report scale (p-value 0.03, adjusted for age and ethnic background). Women reported an earlier mean age of onset of about 2.5 years than men. When adjusted for age and ethnic background a statistical trend remained. No gender differences were found in reported number of MDD episodes. Women revealed more prior suicide attempts than men, whereas suicidal thoughts during the month before assessment were equally reported in both sexes.

Table 2.3. Depressive subtype and DSM-IV comorbidity according to gender in 1,131 outpatients with current MDD

	Men (n=395; 34.9%)	Women (n=736; 65.1%)	p-value	Adjusted p-value ^a
Depression subtype				
Melancholic subtype (n, %)	179 (45.3)	335 (45.5)	1.0	0.59
Atypical subtype (n, %)	92 (23.3)	279 (37.9)	<0.001	0.001
Comorbid DSM-IV-TR disorder (n, %)				
Any anxiety disorder	170 (43.0)	321 (43.6)	0.90	0.61
Posttraumatic stress disorder	43 (10.9)	128 (17.4)	0.004	0.009
Panic disorder	39 (9.9)	79 (10.7)	0.68	0.95
Generalised anxiety disorder	22 (5.6)	27 (3.7)	0.17	0.15
Social phobia	61 (15.4)	76 (10.3)	0.01	0.002
Obsessive compulsive disorder	20 (5.1)	36 (4.9)	0.89	0.76
Any somatoform disorder	50 (12.7)	86 (11.7)	0.63	0.71
Both somatoform and anxiety disorder	20 (5.1)	36 (4.9)	0.89	0.77
Alcohol abuse or dependence	48 (12.2)	30 (4.1)	<0.001	<0.001
Drug abuse or dependence	29 (7.3)	23 (3.1)	0.002	<0.001
Bulimia nervosa	1 (0.3)	16 (2.2)	0.009	0.06
ADHD	11 (2.8)	5 (0.7)	0.007	0.003
Number of comorbid DSM-IV-TR axis I				
0	171 (43.3)	365 (49.6)		
1	149 (37.7)	256 (34.8)		
2	54 (13.7)	90 (12.2)		
>2	21 (5.3)	25 (3.4)		

Abbreviations: MDD: Major Depressive Disorder ; DSM-IV-TR: Diagnostic and Statistical Manuals of Mental disorders, fourth edition, text revision; ADHD: Attention Deficit Hyperactivity Disorder.

^a adjusted for age, ethnic background and MADRS score.

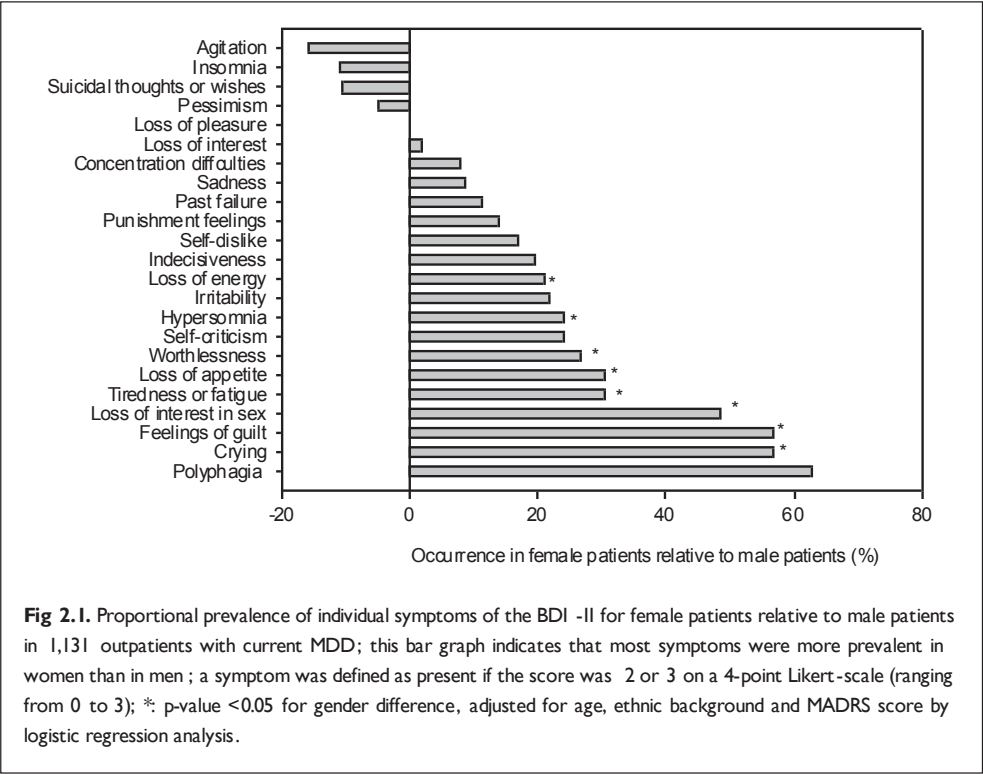
Depressive subtype and DSM-IV comorbidity (Table 2.3)

Men and women equally reported melancholic features, providing the diagnosis 'MDD melancholic subtype' in 45% of the subjects. Atypical features, however, were 63% more prevalent in women than in men (p=0.001 adjusted for age, ethnic background and MADRS score). Significant differences between the sexes were also found in the DSM-IV comorbidity patterns: posttraumatic stress disorder and bulimia nervosa were more

common in women with MDD, whereas social phobia, abuse or dependence on alcohol or drugs, and attention deficit hyperactivity disorder (ADHD) were more common in men. Women were more likely to have a singular depression, i.e. without any comorbid DSM-IV axis I diagnosis.

Symptoms of depression

We found several differences in reported symptoms between men and women, as measured with the BDI-II self-report scale. After adjustment for age, ethnic background and MADRS score, women reported significantly more loss of energy, hypersomnia, worthlessness, loss of appetite, tiredness, loss of interest in sexual activity, feelings of guilt and crying than men. In figure 2.1, the proportional differences in occurrence of BDI-II symptoms between women and men are shown.



Generic health status (Table 2.4)

After adjusting for age and ethnic background, self-reported generic health status was significantly lower in women than in men on six of eight subscales of the SF-36: physical functioning, social functioning, physical problems, emotional problems, vitality and bodily pain. No gender-specific differences were found in the subscales mental health and general health.

Table 2.4. Generic health status: SF-36 subscale scores according to gender in 1,131 outpatients with current MDD

SF-36 subscale (mean ± SD)	Men (n=444; 35.4%)	Women (n=809; 64.6%)	p-value	Adjusted p-value ^a
Physical functioning	72.3 ± 24.7	69.3 ± 25.1	0.06	0.002
Social functioning	38.8 ± 24.0	35.4 ± 23.9	0.02	0.01
Physical problems	32.8 ± 37.6	29.4 ± 35.6	0.14	0.02
Emotional problems	22.3 ± 30.5	17.1 ± 28.5	0.005	0.002
Mental health	33.2 ± 14.6	33.0 ± 15.3	0.84	0.65
Vitality	29.0 ± 15.5	26.7 ± 14.8	0.02	0.009
Bodily pain	65.3 ± 27.6	60.0 ± 28.4	0.002	<0.001
General health	48.8 ± 21.3	47.7 ± 20.3	0.37	0.26

Abbreviations: MDD: Major Depressive disorder ; SF-36: Short Form 36 (RAND 36); Scores are on a 100 point-scale. A higher score corresponds to better functioning.
^a adjusted for age and ethnic background.

Discussion

We found important differences in MDD severity and symptom profile between men and women in our large secondary and tertiary psychiatric care naturalistic sample of MDD-outpatients. Women reported a significantly more severe phenotype, with more severe symptoms of depression as compared to men on several self-report scales (i.e., BDI-II, BSI and SF-36) and more symptoms on the BDI-II, but no differences in symptom severity were observed on most observer-rated scales (i.e., MADRS and BAS). The second major finding was that more women than men suffered from atypical features of depression and comorbid PTSD, whereas men suffered more from psychomotor inhibition, comorbid social phobia, ADHD and alcohol or drug abuse than women.

Since our patient sample reflects a naturalistic treatment seeking population in psychiatric specialty care, our findings add importantly to the existing literature on gender differences in MDD, which consists mainly of information based on data from RCTs and community samples. It is known that RCTs are much more prone to selection-bias than naturalistic patient samples, partly due to the strict selection-criteria (e.g., patients

with suicidal tendencies or comorbidity are generally excluded) that are often used in RCTs. Such patient selection criteria limit the generalisability of findings to real-world clinical practice (Zetin, 2007; Zimmerman, 2002; Zimmerman, 2005). Specifically, patients who would be excluded from an antidepressant efficacy trial may be more likely to have somatic and psychiatric comorbidity, more previous episodes of MDD, greater psychosocial impairment and more personality pathology (Zimmerman, 2005). In addition, community samples may also poorly reflect a treatment seeking naturalistic sample.

Although the STAR*D study more closely reflected a real-life treatment-seeking patient population because less strict inclusion criteria were used compared to most RCTs, a selection has still been made from patients who chose and gave informed consent to participate in the trial.

Our main finding that women tend to report more and more severe symptoms than men is largely consistent with previous studies in non-naturalistic patient samples (Frank et al., 2005; Kornstein et al., 2000; Marcus et al., 2005; Marcus et al., 2008; Verhagen, 2008). Similar scores for men and women on clinician rated scales of depression (Hamilton Depression Rating Scale (HAMD), MADRS) were found in most studies (Carter et al., 2000; Frank et al., 1988; Kornstein et al., 2000; Marcus et al., 2005), with the exception of the 2008 STAR*D replication study, which found a slightly higher HAMD and clinician-rated Inventory of Depressive Symptomatology (IDS) score in women versus men (Marcus et al., 2008). Higher scores on the BDI and SF-36 self-report scales in women with MDD as compared to men were also reported by Kornstein et al (Kornstein et al., 2000). In general, self-report scales tend to show a more subjective score as compared to clinician-rated scales, although the latter scales may suffer from specific forms of rater-bias. These errors may be compensated for by combining the two types of rating-scales (Möller, 2009). Our findings also suggest that subjective distress, and not objective distress, measured in both the amount of symptoms, symptom severity and health status, is higher in depressed women than in depressed men. Alternatively, men could be less inclined to admit depressive symptoms and its impact on self-report scales. The fact that the research nurses were all female may also have played a role, although in the latter case one would expect to find differences in observational scales as well. Young et al. did not find any gender-related rater bias in a comparable setting (Young et al., 1990).

Not only did we find more subjective distress in women than men, we also showed that their subjective symptom profile was different (see figure 2.1). These quantitative and qualitative differences may have several implications for diagnosis and treatment. Firstly, if in studies or treatment settings depressed patients are included based on a cut-off score on an observer-rated scale, women with scores just below the cut-off may be

withhold treatment they need in case their subjective scores would be taken into account. How large this group is and to what extent these patients indeed would be able to profit from treatments should be a subject of further investigations. Secondly, the differences in subjective symptom profile suggest that psychotherapeutic interventions could be more effective if these gender differences were taken into account. The same is true for pharmacological treatment as pharmacological treatment choices are often based on clinical characteristics (Trivedi et al., 2008; Zimmerman et al., 2004). Thirdly, as subjective differences in symptoms are related to gender, these differences might also be related to the differences in hormonal status between women and men. This also needs further study.

Other findings in both clinical trial and population-based samples that are in line with our findings include the age of onset of depression, suicidality and rates of atypical depression. We found that women reported an earlier onset of depression, in line with previous studies (Kessler, 2003; Marcus et al., 2005). However, this younger age of onset was confounded by the age at assessment – lower in women than in men – and ethnic background. The 30% higher proportion of female patients than male patients with prior suicide attempts in our study is consistent with most studies (Marcus et al., 2005; Marcus et al., 2008; Petronis, 1990; Rapaport, 1995). The high percentage can be explained by the fact that the sample consisted of secondary and tertiary care patients, which implicates rather severe psychopathology. In most RCTs, suicidality is an exclusion criterion. Hence, reported information about suicidality in RCTs is likely to be an underestimate of the true suicidality in MDD. We found a higher predominance of atypical features of depression in women than in men. A higher female preponderance of atypical depression is generally found in the literature, although different definitions or constructs of atypical depression are used (Angst et al., 2006; Khan et al., 2002; Silverstein, 2002; Smith et al., 2008). This higher level of atypical depression in women could reflect a different underlying biological aetiology for depression or a different expression of a similar underlying aetiology in men and women (Gold, 2002; Smith et al., 2008). Comorbidity patterns that differed between the sexes, eg. more ADHD and alcohol and drug abuse in men, and more bulimia and PTSD in women, were previously found in population-based and clinical trial studies (Breslau, Schultz, & Peterson, 1995; Kornstein & Sloan, 2005; Marcus et al., 2008). But interestingly, overall we found no higher number of comorbid anxiety disorders in women, in contrast to most studies (Breslau, Schultz, & Peterson, 1995; Kornstein & Sloan, 2005; Marcus et al., 2005; Marcus et al., 2008)

Several theories have been proposed to explain gender differences in MDD, for example differences in treatment seeking behaviour between men and women, psychological factors, genetic factors, neuroendocrine factors like differences in sex hormones, and changes during the menstrual cycle in premenopausal women. Exact mechanisms, however, remain unclear (Kendler, 2001; Kuehner, 2003; Middeldorp et al., 2006; Piccinelli & Wilkinson, 2000).

Strengths and limitations

Strengths of this study are the large study population of over 1,100 MDD patients, the routinely, as part of the normal clinical process adapted computerised assessments by specially trained research nurses and the fact that this population reflects a naturalistic treatment-seeking population in psychiatric specialty care.

Our study also has several potential limitations. Firstly, no information about somatic comorbidity, familial status or use of medication was ascertained. A previous population-based study showed evidence for gender-related differences in comorbidity, dependent on familial status of MDD (Verhagen, 2008). The fact that women reported more bodily pain or more impaired functioning due to pain on the SF-36 (table 4) could correlate with the presence of more somatic comorbidity in women. Secondly, we were not able to assess all referred MDD patients, because about 20% of all patients in RD and LUMC did not have a baseline ROM assessment due to language problems or logistical reasons. So in spite of the fact that only a minimal selection was made in our patients due to the routinely applied ROM assessments, this sample may not be wholly representative for all MDD outpatients. Thirdly, since DSM-IV atypical features have not been assessed systematically, we defined atypical depression by the presence of two or more from the BDI-II items of hyperphagia, hypersomnia and fatigue (Angst et al, 2006) This construct is not per se comparable with DSM-IV atypical depression since information about interpersonal rejection sensitivity and mood reactivity was lacking.

We conclude that in our naturalistic MDD-sample, depression severity in women was higher compared to men when measured on self-report scales (BDI-II, BSI), whereas no difference was found on observational scales (MADRS, BAS). In addition, reported health status was worse in women than in men, measured on the SF-36. These findings may be an explanation for the higher predominance of women with MDD that attend psychiatric specialty care. The prognostic and therapeutic implications of these findings should be further studied. Furthermore, the need for studying psychological factors, genetic liability and biological mechanisms underlying the gender differences in depression and its phenomenology is of importance.

References

- Aaronson, N.K., 1998. Translation, validation, and norming of the Dutch Language version of the SF-36 Health Survey in community and chronic disease populations. *Journal of Clinical Epidemiology* 51, 1055.
- Angst, J., Gamma, A., Benazzi, F., Silverstein, B., Ajdacic-Gross, V., Eich, D., Rössler, W., 2006. Atypical depressive syndromes in varying definitions. *European Archives of Psychiatry and Clinical Neuroscience* 256, 44-54.
- Angst, J., Gamma, A., Gastpar, M., Lepine, J. P., Mendlewicz, J., Tylee, A., 2002. Gender differences in depression. Epidemiological findings from the European DEPRES I and II studies. *European Archives of Psychiatry and Clinical Neuroscience* 252, 201-209.
- Beck, A.T., Steer, R.A., Brown, O.K. 1996, *Beck Depression Inventory Manual*, 2nd ed., Psychological Corporation, San Antonio, TX.
- Beck, A.T., Steer, R.A., Carbin, M.G., 1988. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review* 8, 77-100.
- de Beurs, E. & Zitman, F.G., 2006. The Brief Symptom Inventory (BSI): Reliability and validity of a practical alternative to SCL-90. *Maandblad Geestelijke Volksgezondheid* 61, 120-141.
- Breslau, N., Schultz, L., Peterson, E., 1995. Sex-differences in depression - A role for preexisting anxiety. *Psychiatry Research* 58, 1-12.
- Carter, J.D., Joyce, P.R., Mulder, R.T., Luty, S.E., McKenzie, J., 2000. Gender differences in the presentation of depressed outpatients: a comparison of descriptive variables. *Journal of Affective Disorders* 61, 59-67.
- Derogatis, L.R., 1977. Confirmation of the dimensional structure of the SCL-90: A study in construct validation. *Journal of Clinical Psychology* 33, 981-989.
- Derogatis, L.R. & Melisaratos, N., 1983. The Brief Symptom Inventory: an introductory report. *Psychological Medicine* 13, 595-605.
- Fava, M., Abraham, M., Alpert, J., Nierenberg, A.A., Pava, J.A., Rosenbaum, J.F., 1996. Gender differences in Axis I comorbidity among depressed outpatients. *Journal of Affective Disorders* 38, 129-133.
- Frank, E., Carpenter, L.L., Kupfer, D.J., 1988. Sex-Differences in Recurrent Depression - Are there any that are significant. *The American Journal of Psychiatry* 145, 41-45.
- Goekoop, J.G., 1992. A factor-analytic study using the Comprehensive Psychopathological Rating Scale. *Acta Psychiatrica Scandinavica* 86, 306-312.
- Gold, P.W., 2002. Organisation of the stress system and its dysregulation in melancholic and atypical depression: high vs. low CRH/NE states. *Molecular Psychiatry* 7, 254-275.
- Grigoriadis, S., Konarski, J.Z., Kennedy, S.H., Mancini, D.A., McIntyre, R.S., 2007. Sex differences in antidepressant response in a Canadian primary-care sample. *Journal of Clinical Psychopharmacology* 27, 95-98.
- Hildebrandt, M.G., 2003. Gender and depression: a study of severity and symptomatology of depressive disorders (ICD-10) in general practice. *Acta Psychiatrica Scandinavica* 107, 197-202.
- Kendler, K.S., 2001. Genetic risk factors for major depression in men and women: similar or different heritabilities and same or partly distinct genes? *Psychological medicine* 31, 605-616.

- Kessler, R.C., 2003. Epidemiology of women and depression. *Journal of Affective Disorders* 74, 5-13.
- Kessler, R.C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K.R., Rush, A.J., Walters, E.E., Wang, P.S., 2003. The epidemiology of major depressive disorder - Results from the National Comorbidity Survey Replication (NCS-R). *Journal of the American Medical Association* 289, 3095-3105.
- Kessler, R.C., Chiu, W.T., Demler, O., Walters, E.E., 2005. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry* 62, 617-627.
- Khan, A.A., Gardner, C.O., Prescott, C.A., Kendler, K.S., 2002. Gender differences in the symptoms of major depression in opposite-sex dizygotic twin pairs. *The American Journal of Psychiatry* 159, 1427-1429.
- Kornstein, S.G. & Sloan D.M., 2005 "Depression and Gender." In: *The American Psychiatric Publishing Textbook of Mood disorders*, Stein D.J., Kupfer D.J., Schatzberg A.F., (eds.), American Psychiatric Publishing, Washington, DC.
- Kornstein, S.G., Schatzberg, A.F., Thase, M.E., Yonkers, K.A., McCullough, J.P., Keitner, G.I., Gelenberg, A.J., Ryan, C.E., Hess, A.L., Harrison, W., Davis, S.M., Keller, M.B., 2000. Gender differences in chronic major and double depression. *Journal of Affective Disorders* 60, 1-11.
- Kuehner, C., 2003. Gender differences in unipolar depression: an update of epidemiological findings and possible explanations. *Acta Psychiatrica Scandinavica* 108, 163-174.
- Leclercq, Y., 1997. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *European Psychiatry* 12, 224-231.
- Marcus, S.M., Kerber, K.B., Rush, A.J., Wisniewski, S.R., Nierenberg, A., Balasubramani, G.K., Ritz, L., Kornstein, S., Young, E.A., Trivedi, M.H., 2008. Sex differences in depression symptoms in treatment-seeking adults: confirmatory analyses from the Sequenced Treatment Alternatives to Relieve Depression study. *Comprehensive Psychiatry* 49, 238-246.
- Marcus, S.M., Young, E.A., Kerber, K.B., Kornstein, S., Farabaugh, A.H., Mitchell, J., Wisniewski, S.R., Balasubramani, G.K., Trivedi, M.H., Rush, A.J., 2005. Gender differences in depression: findings from the STAR*D study. *Journal of Affective Disorders* 87, 141-150.
- Middeldorp, C.M., Wray, N.R., Andrews, G., Martin, N.G., Boomsma, D.I., 2006. Sex differences in symptoms of depression in unrelated individuals and opposite-sex twin and sibling pairs. *Twin Research and Human Genetics* 9, 632-636.
- Möller, H.J., 2009. Standardised rating scales in Psychiatry: Methodological basis, their possibilities and limitations and descriptions of important rating scales. *The World Journal of Biological Psychiatry* 10, 6-26.
- Montgomery, S.A., 1979. A new depression scale designed to be sensitive to change. *The British Journal of Psychiatry* 134, 382-389.
- Muller, M.J., 2003. Differentiating moderate and severe depression using the Montgomery-Åsberg depression rating scale (MADRS). *Journal of Affective Disorders* 77, 255-260.
- Nolen-Hoeksema, S., 1987. Sex differences in unipolar depression: evidence and theory. *Psychological Bulletin* 101, 259-282.
- Petronis, K. R., 1990. An epidemiologic investigation of potential risk factors for suicide attempts. *Social Psychiatry and Psychiatric Epidemiology* 25, 193-199.
- Piccinelli, M. & Wilkinson, G., 2000. Gender differences in depression. Critical review. *The British Journal of Psychiatry* 177, 486-492.

- Rapaport, M.H., 1995. Gender differences in outpatient research subjects with affective disorders: a comparison of descriptive variables. *The Journal of Clinical Psychiatry* 56, 67-72.
- Rush, A.J., Fava, M., Wisniewski, S.R., Lavori, P.W., Trivedi, M.H., Sackeim, H.A., Thase, M.E., Nierenberg, A.A., Quitkin, F.M., Kashner, T.M., Kupfer, D.J., Rosenbaum, J.F., Alpert, J., Stewart, J.W., McGrath, P.J., Biggs, M.M., Shores-Wilson, K., Lebowitz, B.D., Ritz, L., Niederehe, G., 2004. Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design *Controlled Clinical Trials* 25, 119-142.
- Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry* 59 Suppl. 20, 22-33.
- Silverstein, B., 2002. Gender differences in the prevalence of somatic versus pure depression: a replication. *The American Journal of Psychiatry* 159, 1051-1052.
- Smith, D.J., Kyle, S., Forty, L., Cooper, C., Walters, J., Russell, E., Caesar, S., Farmer, A., McGuffin, P., Jones, I., Jones, L., & Craddock, N., 2008. Differences in depressive symptom profile between males and females. *Journal of Affective Disorders* 108, 279-284.
- Snaith, R.P., 1986. Grade scores of the Montgomery-Åsberg Depression and the Clinical Anxiety Scales. *The British Journal of Psychiatry* 148, 599-601.
- Trivedi, M.H., Hollander, E., Nutt, D., Blier, P., 2008. Clinical evidence and potential neurobiological underpinnings of unresolved symptoms of depression. *Journal of Clinical Psychiatry* 69, 246-258.
- 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 & Psychiatry* 47, 970-975.
- Verhagen, M., 2008. Familiality of major depressive disorder and gender differences in comorbidity. *Acta Psychiatrica Scandinavica* 118, 130-138.
- van Vliet, I.M. & de Beurs, E., 2007. [The MINI-International Neuropsychiatric Interview. A brief structured diagnostic psychiatric interview for DSM-IV en ICD-10 psychiatric disorders]. *Tijdschrift voor Psychiatrie* 49, 393-397.
- Ware Jr, J.E., 1992. The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection. *Medical Care* 30, 473-483.
- World Health Organisation 1990, Composite International Diagnostic Interview (CIDI), version 1.1, WHO edn., Geneva.
- Young, M.A., Scheftner, W.A., Fawcett, J., Klerman, G.L., 1990. Gender differences in the clinical features of unipolar major depressive disorder. *Journal of Nervous and Mental Disorders* 178, 200-203.
- Zetin, M., 2007. Relevance of exclusion criteria in antidepressant clinical trials: a replication study. *Journal of Clinical Psychopharmacology* 27, 295-301.
- Zimmerman, M., Posternak, M., Friedman, M., Attiullah, N., Baymiller, S., Boland, R., Berlowitz, S., Rahman, S., Uy, K., Singer, S., 2004. Which factors influence psychiatrists' selection of antidepressants? *The American Journal of Psychiatry* 161, 1285-1289.
- Zimmerman, M., 2002. Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? *The American Journal of Psychiatry* 159, 469-473.

Zimmerman, M. Chelwinski, I., Posternak., M.A., 2005. Generalisability of antidepressant efficacy trials: differences between depressed psychiatric outpatients who would or would not qualify for an efficacy trial. *The American Journal of Psychiatry* 162, 1370-1372.

Chapter 3

Pre-adult versus adult onset major depressive disorder in a naturalistic patient sample: the Leiden Routine Outcome Monitoring Study

Martijn van Noorden
Sanne Minkenberg
Erik Giltay
Margien den Hollander-Gijsman
Yanda van Rood
Nic van der Wee
Frans Zitman

Psychological Medicine (2010), 41, 1407-1417

Abstract

Background: Pre-adult onset of Major Depressive Disorder (MDD) may predict a more severe phenotype of depression. As data from naturalistic psychiatric specialty care settings are scarce, we examined phenotypic differences between pre-adult and adult onset MDD in a large sample of consecutive outpatients.

Methods: Altogether, 1,552 outpatients, mean age 39.2 ± 11.6 years, were diagnosed with current MDD on the Mini International Neuropsychiatric Interview-Plus (MINI-Plus) as part of the usual diagnostic procedure. A total of 1,105 patients (71.2%) had complete data on all variables of interest. Pre-adult onset of MDD was defined as having experienced the signs and symptoms of a first Major Depressive Episode before the age of 18 years. Patients were stratified according to the age at interview (20-40 / 40-65 years). Correlates of pre-adult onset were analysed using logistic regression models adjusted for age, age squared and gender.

Results: Univariable analyses showed that pre-adult onset of MDD had a distinct set of demographic (e.g. less frequently living alone) and clinical correlates (more comorbid DSM-IV-Text Revision diagnoses, more social phobia, more suicidality). In the multivariable model we found an independent association only for a history of suicide attempts (OR 3.15; 95% CI: 1.97-5.05) and current suicidal thoughts (OR 1.81; 95% CI 1.26-2.60) in patients with pre-adult versus adult onset MDD.

Discussion: Pre-adult onset of MDD is associated with more suicidality than adult onset MDD. Age of onset of depression is an easy to ascertain characteristic that may help clinicians in weighing suicide risk.

Introduction

Several studies suggest that pre-adult and adult onset major depressive disorder (MDD) are two distinct forms of MDD in terms of pathophysiology and phenomenology (see review by Kaufman et al., 2001). Diagnostic criteria according to the Diagnostic and Statistical Manual of Mental Disorders-Text Revision (DSM-IV-TR) for MDD are independent of the age of onset of the first major depressive episode (MDE; American Psychiatric Association, 2000). This contrasts with dysthymic disorder, for which the DSM-IV-TR makes a distinction in age of onset before or after 21 years of age (early onset and late onset, respectively).

In previous studies, demographic characteristics of pre-adult versus adult onset MDD were female gender, being unmarried (Benazzi, 2000; Zisook et al., 2004; Zisook et al., 2007a), impaired social and occupational functioning (Zisook et al., 2007a), and lower education (Parker et al., 2003). The course of the depressive disorder also differed significantly in some studies, with pre-adult onset MDD being associated with more depressive episodes (Benazzi, 2000; Klein, 1999; Ramklint & Ekselius, 2003; Zisook et al., 2004; Zisook et al., 2007a) and more chronicity (Benazzi, 2000; Coryell et al., 2009; Klein, 1999; Parker et al., 2003; Zisook et al., 2004; Zisook et al., 2007a; Zisook et al., 2007b). In addition, pre-adult MDD patients showed more suicidality (Benazzi, 2000; Zisook et al., 2004; Zisook et al., 2007a; Zisook et al., 2007b; Thompson, 2008), a higher amount of medical comorbidity (Zisook et al., 2007a), more familiarity of depression (Klein, 1999; Parker et al., 2003; Kendler et al., 2005; Zisook et al., 2007b), and more alcohol and drug abuse (Klein, 1999; Parker et al., 2003; Zisook et al., 2007b). Early childhood risk factors like motor skill deficits, perinatal insults and caretaker instability, criminality and psychopathology in the families of origin were associated with pre-adult onset MDD in a birth cohort that had been followed from childhood up to age 26 (Jaffee et al., 2002). Taken together, these data suggest that pre-adult onset MDD points towards a more severe form of illness.

Previous investigations have been undertaken in various study populations, using varying study designs. Prospective findings from a New Zealand birth-cohort suggested that both heritable and childhood psychosocial factors contributed to a pre-adult onset of depression (Jaffee et al., 2002). These findings were consistent with results from family studies which suggested that pre-adult onset MDD might be more strongly associated with genetic factors and early childhood psychosocial risk factors (Kovacs et al., 1997; Neuman et al., 1997; Klein et al., 2001). Several studies reported a different course of illness of pre-adult onset and adult onset MDD. Adolescent MDD has been associated with elevated rates of subsequent MDEs in early adulthood in prospective case control studies (Harrington et al., 1990; Weissman et al., 1999). Yet, most studies used cross-

sectional designs in selected populations. In three analyses in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) Study population (Zisook et al., 2004; Zisook et al., 2007a; Zisook et al., 2007b), using data from over 4,000 selected MDD patients the most prominent findings in pre-adult onset MDD patients were more psychiatric comorbidity and more suicidality. These findings were largely consistent with the findings in a large US community sample of 9,282 people (Thompson, 2008).

It is well established that clinical treatment samples may be prone to selection bias, whereas population-based samples may not reflect a treatment seeking population either (Zimmerman et al., 2002). Since data on naturalistic patients are generally considered to represent 'real life patients' more closely than selected populations, naturalistic data on differences between pre-adult onset versus adult onset MDD would be of great value. Therefore, the aim of the present study was to evaluate differences in demographic correlates, comorbidity, symptomatology and general health status between pre-adult and adult onset MDD in a large naturalistic outpatient sample from psychiatric specialty care. We used the Leiden Routine Outcome Monitoring study baseline sample (de Beurs et al., 2011; van Noorden et al., 2010). We hypothesised that pre-adult onset MDD would be characterised by markers of greater severity as compared with adult onset MDD. Pre-adult onset of MDD was defined as first MDE before the age of 18 years.

Methods

Study design

The Leiden Routine Outcome Monitoring study baseline cohort sample comprised 3,798 adult outpatients who were referred for treatment of a mood, anxiety, or somatoform (MAS) disorder to the Dutch Regional Mental Health Provider (RMHP) Rivierduinen (RD) or the psychiatric outpatient department of the Leiden University Medical Center (LUMC) in the Western part of the Netherlands, between January 2004 until December 2006 (van Noorden et al., 2010; de Beurs et al., 2011). At baseline, subjects were assessed as part of the Routine Outcome Monitoring (ROM) procedure. In ROM, all patients referred to RD or LUMC for treatment of a MAS disorder are routinely assessed, at baseline and at fixed intervals during treatment, with an extensive battery of psychometric instruments administered by specially trained research nurses. ROM is a method for the systematic collection of data on the diagnostic status and severity of complaints to assess the effectiveness of treatments in everyday clinical practice. The only exclusion criteria for ROM are insufficient mastery of the Dutch language, and inability to perform the assessment

procedure due to severity of symptoms. In practice, more than 80% of the patients referred to the LUMC or RD for treatment of a MAS disorder are enrolled in ROM. During the first session, psychopathology was assessed with the Dutch translation of the Mini International Neuropsychiatric Interview-Plus 5.0.0.-R, structured diagnostic interview (MINI-Plus) developed to assess the presence of Axis-I disorders according to the DSM-IV-TR diagnostic criteria (Sheehan et al., 1998; van Vliet & de Beurs, 2007). The MINI-Plus has been validated with the Composite International Diagnostic Interview (CIDI; World Health Organisation, 1990) diagnoses and has good psychometric properties (Lecrubier et al., 1997). Patient data were stored anonymously in the Psychiatric Academic Registration Leiden (PAREL) database and were accessible for research purposes only. This procedure has been approved by the Medical Ethical Committee of the LUMC. For the present study, only the baseline ROM data were used.

Study population

In the present analyses all patients with a current DSM-IV-TR MDD diagnosis according to the MINI-Plus (age range 20-65) were included. We used only baseline ROM assessments, administered at intake. No distinction was made between MDD as principal or comorbid diagnosis. Concomitant current DSM-IV-TR disorders were recorded as well. The presence of DSM-IV-TR melancholic features of depression was systematically assessed with the MINI-Plus, whenever the DSM-IV-TR criteria for MDD were fulfilled. Patients with a bipolar disorder or lifetime psychotic illness were excluded, but a diagnosis of psychotic depression was allowed. Only patients with complete data on all variables of interest were included in the initial analyses. 1,552 (40.9%) of 3,798 patients fulfilled the DSM-IV-TR criteria of MDD, current episode on the MINI-Plus, and were aged 20-65. Of these 1,552 MDD patients, 1105 (71.2%) had complete data on all variables of interest and were included in the present analyses. The excluded 447 patients differed statistically significant in age (mean age 40.6, vs. 39.3 for included patients, respectively; $t=2.06$, $p=0.04$), age of onset (mean age 31.4 vs. 28.3 for included patients, respectively, $t=4.21$, $p<0.001$), MADRS score (mean score 24.2 vs. 25.5 for included patients, respectively; $t=-2.54$, $p=0.02$), but not in gender, total BSI score, current suicidal thoughts or history of suicide attempts.

Age of onset

Age of onset of MDD was defined as the age at which the first MDE initiated, regardless of whether treatment was sought. In the MINI-Plus, when a current MDD was diagnosed, the age of onset of the first episode was assessed by the question: 'How old were you when you first experienced these symptoms of depression, for at least two weeks?' In our study

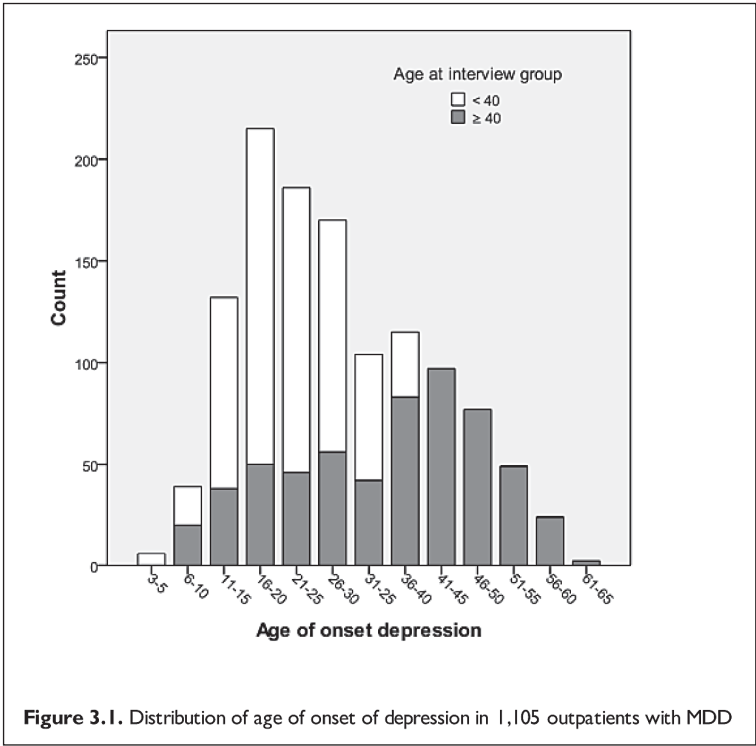
we defined pre-adult onset of MDD as an onset of MDD before age 18, in accordance with the existing literature on this subject. Various cut-off points to distinguish between pre-adult and adult onset MDD have been used in previous studies. Cut-off ages varied from 17 (Jaffee et al., 2002), 18 (Alpert et al., 1999; Benazzi, 2000; Fava et al., 1996; Zisook et al., 2004; Zisook et al., 2007a; Zisook et al., 2007b), 21 (Klein, 1999), 25 (Parker et al., 2003) and 26 (Ramklint and Ekselius, 2003) to 30 (Thompson, 2008), and despite the lack of widely accepted consensus most authors use a cut-off age of 18 years. In sensitivity analyses we re-analysed our data with a cut-off age of 25 years (see statistical analyses).

Other variables

Demographic variables were obtained using a self-report questionnaire that assessed ethnic background, marital status, housing situation, educational status, and employment status. A Dutch ethnic background was assumed when the patient and both parents were born in the Netherlands.

Depression severity was assessed with the Montgomery-Åsberg Depression Rating Scale (MADRS; Montgomery, 1979), a 10-item observer-rated scale that measures symptoms of depression on a 7-point Likert-scale. The MADRS has a good internal consistency and reliability (Montgomery, 1979). In addition to the continuous measure, three categories were distinguished; a MADRS score of <20 for mild depression, 20-35 for moderate depression, and ≥ 35 for severe depression (Muller, 2003; Snaith, 1986). Post-hoc, we analysed MADRS item 10, that assesses suicidal thoughts. We dichotomised in symptom absent, if the score was 0 or 1, and symptom present if the score was 2 or higher. The Brief Anxiety Scale (BAS), an observer-rated scale that measures symptoms of anxiety, was used to assess severity of anxiety (Tyrer et al., 1984). General psychopathological symptoms were scored on the Brief Symptom Inventory (BSI). The BSI is a short version of the Symptom Checklist (SCL-90), a self-report instrument that measures psychopathological symptoms in several domains, e.g. somatic symptoms, depressive symptoms, anxiety symptoms and hostility on a 5-points Likert-scale (de Beurs & Zitman, 2006; Derogatis & Melisaratos, 1983). The BSI has shown good internal consistency, reliability and validity (Derogatis, 1977). In addition to the BSI-total score we studied presence of specific symptoms on the depressive and anxiety subscales of the BSI. A score of 0 or 1 on each individual item was defined as the absence of a symptom, whereas a score of 2, 3 or 4 was defined as the presence of a symptom. A history of suicide attempts and current suicidal thoughts were assessed with the MINI-Plus. In section C of the MINI-Plus, the following questions were asked respectively: "Did you ever in your life try to commit suicide?" and "During the past month, did you ever think about suicide?"

Comorbid DSM-IV-TR disorders, e.g. anxiety disorders, somatoform disorders and alcohol or drug related disorders were also assessed with the MINI-Plus. Generic health status was assessed with the Dutch version of the Short-Form-36 health survey (SF-36), a 36-item self-report questionnaire that measures health status in eight domains: physical functioning, role limitations due to physical health, bodily pain, general health perceptions, vitality, social functioning, role limitations due to emotional problems, and mental health (Aaronson et al., 1998; Ware, Jr. & Sherbourne, 1992).



Statistical analyses

All group comparisons were made between pre-adult onset (age<18 years) and adult onset (age≥18 years) MDD-groups. T-tests for independent samples were used for continuous data, while categorical data were analysed with χ^2 tests. In all previous cross-sectional studies, the pre-adult onset MDD patients were younger than adult onset MDD patients at the time of assessment. Despite the fact that many studies adjusted for age, residual confounding by age at interview is possible as several demographic factors are highly associated with age (e.g., marital status, duration of disease, and comorbidity).

Therefore, we stratified the study population into two age strata (20-40 and 40-65 years of age at interview). If both age strata would yield similar associations and odds ratios, the confounding effects by age are likely to be small. Stratification based on the age of 40 years was chosen because this age is generally considered to be the beginning of middle age, and it was close to the median age of the sample (i.e., 38 years). Also, it was assumed that the associations between the clinical characteristics and age of onset would vary by gender. Therefore we also made adjustments for gender and current age by including these characteristics in a logistic regression model with forced entry (adjusted analysis). To account for the potential curvilinear confounding effects of age, a quadratic term (age-squared) was added to the model. In the logistic regression model, we analysed the univariable findings with a $p < 0.1$ in the two age strata and the combined group. Sensitivity analyses were performed in which patients with an age at interview of 20-25 years were excluded from the analyses, to rule out the possibility that any differences in outcome could be explained by the very short disease duration. Sensitivity analyses were subsequently performed on the complete sample of 1,552 MDD patients while using multiple imputations of missing data. Finally, sensitivity analyses were performed in which a cut-off age of 25 instead of 18 years was used to distinguish between pre-adult onset and adult onset MDD. The statistical significance was set at $p < 0.05$. When appropriate, Bonferroni correction for multiple testing was applied. Statistical analysis was carried out using the Statistical Package for the Social Sciences (SPSS) 17.0 for Windows.

Results

Sociodemographic characteristics

The sample consisted of 711 women and 394 men (35.7%). The mean age of the subjects was 39.2 (SD 11.6) years (range 20-65). Figure 3.1 illustrates the distributions of age of onset of MDD and age at interview of the 1,105 patients. In table 3.1 the sociodemographic characteristics of the study population are presented. In total, 246 patients (22.3%) had an early-onset form of MDD. These patients were younger than adult onset patients at the time of interview in both strata (in 20-40 group: median age 25 vs. 31 years, respectively [$p < 0.001$], in 40-65 group: median age 46 vs. 49 years [$p = 0.03$]). No difference in ethnic background was found between pre-adult and adult onset groups.

Table 3.1. Sociodemographic characteristics according to age of onset in 1,105 outpatients aged between 20 and 65 with MDD

	Age 20-40 years			Age 40-65 years		
	Pre-adult onset <18 (n =172)	Adult onset ≥18 (n =410)	p-value	Pre-adult onset <18 (n =74)	Adult onset ≥18 (n =449)	p-value
Female gender (n, %)	119 (69.2)	295 (72.0)	0.50	49 (66.2)	248 (55.2)	0.08
Age at interview (median, p25-p75)	26 (22-32)	31 (26-36)	<0.001 ^a	46 (43-53)	49 (44-55)	0.03 ^a
Ethnic background (n, %)			0.27			0.40
Dutch	134 (77.9)	293 (71.5)		60 (81.1)	373 (83.1)	
Other	38 (22.1)	117 (28.5)		14 (18.9)	76 (16.9)	
Marital status (n, %)			0.006			0.83
Married	62 (36.0)	187 (45.6)		44 (59.5)	250 (55.7)	
Divorced/ widowed	10 (5.8)	42 (10.2)		21 (28.4)	137 (30.5)	
Never married	100 (58.1)	181 (44.1)*		9 (12.1)	62 (13.8)	
Housing situation (n, %)			0.05			0.02
Living alone	59 (34.3)	106 (25.9)		11 (14.9)	129 (28.7)*	
Living with partner	63 (36.6)	191 (46.5)		45 (60.8)	250 (55.7)	
Living with family	50 (29.1)	113 (27.6)		18 (24.3)	70 (15.6)	
Educational status (n, %)			0.28			0.85
Lower education	59 (34.3)	160 (39.0)		40 (54.1)	248 (55.2)	
Higher education	113 (65.7)	250 (61.0)		34 (45.9)	201 (44.8)	
Employment situation (n, %)			0.05			0.05
Working full-time	36 (20.9)	82 (20.0)		8 (10.8)	78 (17.4)	
Working part time	27 (15.7)	83 (20.2)		19 (25.7)	63 (14.0)	
Retired/unemployed	59 (34.3)	98 (23.9)		21 (28.4)	123 (27.4)	
On sick leave	50 (29.1)	147 (35.9)		26 (35.1)	185 (41.2)	

MDD denotes Major Depressive Disorder ; * Indicates significant difference post-hoc with chi-squared test after Bonferroni correction.

^a Mann-Whitney non-parametric test.

Table 3.2. Rating scale scores and subtype of depression according to age of onset in 1,105 outpatients aged between 20 and 65 with MDD

	Age 20-40 years			Age 40-65 years		
	Pre-adult onset <18 (n =172)	Adult onset ≥18 (n =410)	p-value	Pre-adult onset <18 (n =74)	Adult onset ≥18 (n =449)	p-value
MADRS score (mean ± SD)	24.2 (7.8)	25.0 (7.8)	0.28	26.0 (8.4)	26.1 (7.6)	0.97
MADRS categories (n, %)			0.28			0.65
MADRS <20	44 (25.6)	84 (20.5)		13 (17.6)	81 (18.0)	
MADRS 20-35	117 (68.0)	289 (70.5)		50 (67.6)	318 (70.8)	
MADRS ≥35	11 (6.4)	37 (9.0)		11 (14.9)	50 (11.1)	
BAS score (mean, ± SD)	17.0 (6.3)	17.4 (6.7)	0.59	18.0 (6.5)	17.4 (6.5)	0.43
BSI total score (geometric mean, p5-p95)	1.6 (1-3)	1.4 (0-3)	0.41	1.5 (0-3)	1.4 (0-3)	0.11
Suicidality						
Current suicidal thoughts (n, %)	77 (44.8)	140 (34.1)	0.02	35 (47.3)	153 (34.1)	0.03
History of suicide attempt (n, %)	53 (30.8)	76 (18.5)	0.001	27 (36.5)	87 (19.4)	0.001
MADRS item 10 ≥2 (n, %)	81 (47.1)	155 (37.8)	0.04	34 (45.9)	200 (44.5)	0.82
Depression subtype						
without comorbidity						
(pure depression) (n, %)	74 (43.0)	193 (47.1)	0.37	36 (48.6)	40 (53.5)	0.44
melancholic depression (n, %)	81 (47.1)	166 (40.5)	0.14	41 (55.4)	221 (49.2)	0.32

Abbreviations: MDD: Major Depressive episode ; MADRS: Montgomery-Åsberg Depression Rating Scale; BAS: Brief Anxiety Scale; BSI: Brief Symptom Inventory.

Univariable analyses

Table 3.2 shows the relationship between age at onset and rating scale scores and subtype of depression. The main difference was that patients with pre-adult onset MDD had significantly more often a history of suicide attempts (in 20-40 group 30.8% vs. 18.5% [$p<0.001$]; in 40-65 group 36.5% vs. 19.4% [$p<0.001$]). No differences between the groups were found in MADRS continuous scores, nor in MADRS categories. Only in the 20-40 group, pre-adult onset patients more often had a positive score on the MADRS suicidality item (47.1% vs. 37.8% [$p=0.04$]). No significant differences were found between the early- and adult onset groups on the BSI or BAS scores, nor in the presence of subtypes of depression.

Furthermore, we analysed the differences between pre-adult onset and adult onset groups on the symptom level, by focusing on the anxiety and depression subscales of the BSI. Significant differences in univariable analyses were found for the following two items only: “thoughts of ending your life” and “feeling hopeless about the future”. Consistent with “suicidal thoughts” on the MINI-Plus, these affirmative answers from the depression subscale were more prevalent in patients with pre-adult onset than adult onset MDD (data not shown).

Table 3.3. Current comorbidity according to age of onset in 1,105 outpatients with MDD

	Age 20-40 years			Age 40-65 years		
	Pre-adult onset <18 n =172	Adult onset ≥18 n =410	p-value	Pre-adult onset <18 n =74	Adult onset ≥18 n =449	p-value
Any Anxiety disorder (n, %)	79 (45.9)	165 (40.2)	0.21	28 (37.8)	155 (34.5)	0.58
Posttraumatic stress disorder	34 (19.8)	69 (16.8)	0.40	9 (12.2)	57 (12.7)	0.90
Panic disorder	267 (15.1)	44 (10.7)	0.14	8 (10.8)	40 (8.9)	0.60
Generalised anxiety disorder	5 (2.9)	17 (4.1)	0.47	4 (5.4)	23 (5.1)	0.92
Social phobia	31 (18.0)	45 (11.0)	0.02	12 (16.2)	40 (8.9)	0.05
Obsessive compulsive disorder	12 (7.0)	24 (5.9)	0.61	5 (6.8)	16 (3.6)	0.20
Any Somatoform disorder (n, %)	9 (5.2)	28 (6.8)	0.47	6 (8.1)	36 (8.0)	0.98
Hypochondria	2 (1.2)	8 (2.0)	0.50	2 (2.7)	4 (0.9)	0.18
Body dysmorphic disorder	5 (2.9)	12 (2.9)	0.99	0 (0)	3 (0.7)	0.48
Both somatoform and anxiety disorder (n, %)	10 (5.8)	24 (5.9)	0.99	4 (5.4)	18 (4.0)	0.60
Alcohol dependence / alcohol abuse (n, %)	13 (7.6)	20 (4.9)	0.20	8 (10.8)	39 (8.7)	0.55
Drug dependence / drug abuse (n, %)	14 (8.1)	19 (4.6)	0.09	5 (6.8)	11 (2.4)	0.05
Number of comorbid MAS disorders (n, %)			0.02			0.04
0	65 (37.8)	187 (45.6)		30 (40.5)	240 (53.5)*	
1	62 (36.0)	157 (38.3)		27 (36.5)	149 (33.2)	
≥2	45 (26.2)	66 (16.1)*		17 (23.0)	60 (13.4)	

Abbreviations: MDD: Major Depressive disorder; MAS: Mood, anxiety and somatoform; * Indicates significant difference post-hoc with chi-squared test after Bonferroni correction.

Table 3.3 summarises the presence of comorbid DSM-IV-TR disorders according to age of onset in each group, in univariable analyses. Pre-adult onset individuals had increased rates of comorbid social phobia, but no differences were found in the rates of other DSM-IV-TR anxiety disorders.

Patients with a pre-adult onset MDD were more likely to be dependent on drugs or to abuse drugs and had a higher number of comorbid MAS disorders.

No differences were found for any of the SF-36 subscale scores between pre-adult onset and adult onset MDD patients (data not shown).

Table 3.4. OR for outcomes according to pre-adult age of onset of MDD in 1,105 outpatients aged between 20 and 65 with MDD

	Age 20-40 years		Age 40-65 years		Both groups combined	
	No. of patients	OR for pre-adult onset (95% CI)	No. of patients	OR for pre-adult onset (95% CI)	No. of patients	OR for pre-adult onset (95% CI)
Marital status						
Married	249	1	294	1	543	1
Divorced/widowed	52	0.84 (0.28 – 2.54)	158	1.17 (0.30 – 4.55)	210	1.06 (0.47 – 2.38)
Never married	281	1.01 (0.44 – 2.63)	71	1.26 (0.29-5.44)	352	1.21 (0.57 – 2.59)
Housing situation						
Living alone	165	1	140	1	305	1
Living with partner	254	0.90 (0.35 – 2.28)	295	2.55 (0.59 – 10.92)	549	1.27 (0.58 – 2.75)
Living with family	163	0.58 (0.34 – 0.96)	88	2.51 (1.05 – 6.01)	251	0.93 (0.61 – 1.42)
History of suicide attempts						
No	453	1	409	1	862	1
Yes	129	4.17 (2.25 – 7.72)	114	2.52 (1.12 – 5.67)	243	3.15 (1.97 – 5.05)
Current suicidal thoughts						
No	365	1	335	1	700	1
Yes	217	1.92 (1.22 – 3.01)	188	1.59 (0.84 – 3.01)	405	1.81 (1.26 – 2.60)
Number of comorbid MAS disorders						
0	252	1	270	1	522	1
1	219	0.99 (0.64 – 1.54)	176	1.30 (0.72 – 2.35)	395	1.14 (0.81 – 1.61)
≥2	111	1.58 (0.89 – 2.81)	77	1.20 (0.53 – 2.73)	188	1.55 (0.98 – 2.46)
Comorbid social phobia						
No	506	1	471	1	977	1
Yes	76	1.28 (0.72 – 2.31)	52	1.76 (0.78 -3.95)	128	1.36 (0.85 – 2.17)
Comorbid drug dependence / abuse						
No	549	1	507	1	1056	1
Yes	33	1.22 (0.54 – 2.77)	16	2.63 (0.80 – 8.59)	49	1.59 (0.82 – 3.08)

Abbreviations: OR: Odds Ratios; MDD: Major depressive Disorder ; CI: Confidence Interval; Reference =adult onset MDD

Data adjusted for age at interview and gender with forced entry in logistic regression models. Bold denotes statistical significant difference at $p<0.05$.

Multivariable and sensitivity analyses

The results of the multivariable analyses are presented in table 3.4. After adjusting for age at interview and gender, in both strata pre-adult onset MDD patients significantly more often had a history of suicide attempts compared to adult onset MDD patients (OR in 20-40 group: 4.17; 95% confidence interval [CI]: 2.25-7.72; OR in 40-65 group: 2.52; 95% CI: 1.12-5.67). Furthermore, pre-adult patients in the 20-40 group and in the combined groups more often had current suicidal thoughts. In the 40-65 group, a non-significant

trend was found. In the combined groups, the OR for having a history of suicide attempts in pre-adult onset MDD patients was 3.15 (95% CI: 1.97-5.05), and the OR for having current suicidal thoughts was 1.81 (95% CI: 1.26-2.60). Results of sensitivity analyses as described in the methods section were as follows: In a first sensitivity analysis when excluding all patients between 20 and 25 years, the odds ratio's for a history of suicide attempts and current suicidal thoughts remained largely unchanged (2.75 and 1.66, respectively, for the combined group). In our second sensitivity analysis we aimed to impute missing data and to analyse the complete sample of 1,552 MDD patients in the multivariable model. However, no variables used in the multivariable model were missing, so the analysis could be performed in 1,552 patients without imputation. Odds ratio's for a history of suicide attempts and current suicidal thoughts were 2.71 and 1.52 respectively, for the combined group. In the third sensitivity analysis we used a cut-off age of 25 instead of 18 years to distinguish between pre-adult and adult onset MDD. In this analysis, OR were again largely comparable to the original analysis (2.78 and 1.94, respectively for the combined group).

Discussion

In our naturalistic patient sample we found that patients with pre-adult onset MDD, defined as an onset before age 18, more often had a history of suicide attempts and current suicidal thoughts compared to patients with adult onset MDD. We found no differences in severity of depression measured with the MADRS, nor differences in the presence of comorbid anxiety disorders or differences in self-reported generic health status.

The main finding of more previous suicide attempts and current suicidal thoughts after adjusting for age at interview and gender in multivariable models in patients with pre-adult onset MDD, is in accordance with the results of several studies in different populations and settings. In three studies of Zisook et al. (Zisook et al., 2004; Zisook et al., 2007a; Zisook et al., 2007b), the STAR*D population was used to investigate factors that differentiate early and later onset MDD. The most prominent finding of the first two studies in 1,500 and 2,541 patients, respectively, was a higher rate of suicidality and more previous suicide attempts in the pre-adult onset MDD group (Zisook et al., 2004; Zisook et al., 2007b). In the third study on the combined sample the study population was divided into five age-at-onset groups. Earlier onset was again associated with more suicidality in multivariable analyses adjusted for age at interview, duration of illness and gender (Zisook et al., 2007a). Most other findings of the first study, like more anxiety disorders in the pre-adult onset group, and greater depressive symptom severity could not be replicated in the

second and third study. Our results are also in line with those of the National Comorbidity Survey, a large representative US population study in 9,289 respondents, where a higher degree of suicidal intent was associated with early onset MDD (Thompson, 2008).

In our study, there were no significant differences in a number of clinical characteristics, such as symptom severity or comorbid anxiety disorders, which were found previously in several clinical trial samples to differ between pre-adult and adult onset MDD (Benazzi, 2000; Zisook et al., 2004; Zisook et al., 2007a; Zisook et al., 2007b). For example, after adjusting for age at interview and gender, pre-adult onset was found to be associated with more social and simple phobias and more alcohol abuse or dependence in a sample of 381 adult MDD patients who were recruited for outpatient clinical trials (Alpert et al., 1999). In a study of 269 MDD outpatients, partly consecutive referrals and partly recruited patients, patients with an MDD-onset before age 25 showed more irritability, anxiety, and more alcohol and drug use in multivariable models adjusting for age at interview (Parker et al., 2003). The fact that earlier cross-sectional studies identified other clinical differences between pre-adult onset and adult onset MDD patients might be explained by the differences in study design and populations. These clinical studies often used selected groups of patients that had been recruited for clinical treatment trials. It is known that clinical treatment studies, particularly Randomised Controlled Trials (RCTs), use stringent criteria for patient selection which may reduce the generalisability to routine clinical practice. For example, patients with suicidality, comorbidity or unsuccessful previous treatments are often excluded from RCTs (Zetin & Hoepner, 2007; Zimmerman et al., 2002; Zimmerman et al., 2005). If suicidal patients are excluded, finding an association between suicidality and age of onset would be impossible. (e.g. Alpert et al., 1999; Klein, 1999) On the other hand, one would expect to validate previous findings from more selected populations in a replication study with a naturalistic design. Since the results of studies carried out in naturalistic patient samples are more applicable to 'real-life' patients, the probability of finding 'true effects' is higher in these studies. Indeed, the finding of more suicidality in pre-adult onset MDD was the main finding of the three STAR*D trials, with a design that aimed to have maximum generalisability of patients in both specialty and primary care.

Another remarkable difference between previous cross-sectional studies and our study is the proportion of MDD patients that reported a pre-adult onset of MDD. The percentage (22%) of MDD patients that reported a pre-adult onset MDD in our study is lower than the 35-40% found in other studies (Alpert et al., 1999; Fava et al., 1996; Zisook et al., 2007a) also using a cut-off age of 18. Since both our study and these earlier studies obtained the age of onset retrospectively, this cannot explain the observed differences.

Hence, the differences are probably the result of differences in study design and study samples.

We found no differences between the groups in rates of patients with melancholic depression. The overall rate of 46.1% (509 from 1,105 patients, see table 3.2) is in line with previous prevalence estimates of melancholic depression in psychiatric specialty care outpatients, that vary from 16% to 67% (Mallinckrodt et al., 2005) depending on definitions being used.

The finding that patients with a pre-adult onset MDD are more likely to have attempted suicide and to have current suicidal thoughts could have several explanations. First, the increased occurrence of suicidality in pre-adult onset MDD could point towards a different variant of MDD, with a different genetic and familial load or other childhood experiences. Indeed, previous studies on suicidality in depression found that genetic factors and familial loading were risk factors for suicide attempts in depression as were early traumatic experiences (Malone et al., 1995). Unfortunately, no information about familial status or early trauma was ascertained in ROM in the study period. Second, MDD with an onset before age 18 could reflect a more severe variant of MDD. Suicidality and suicide attempts generally occur in severely depressed patients and hence can be interpreted as a marker of severity of psychopathology (Forman et al., 2004). However, this explanation is not supported by the results of our study as we did not find differences between the pre-adult and late-adult onset MDD patient on global measures of severity, e.g. MADRS scores or BSI total scores. A third explanation for the increased suicidality could be found in the time passed by since the onset of MDD as patients with pre-adult onset of MDD generally have been depressed longer or more often. Because we adjusted for current age and stratified in two age groups, we did not adjust for the highly correlated disease duration variable. Furthermore, in a sensitivity analysis in which we excluded all patients aged 20-25, the odds ratio's remained largely unchanged. Hence, we believe that our data do not support the explanation of the increased suicidality being the result of the longer disease duration or in patients with pre-adult onset MDD. A final explanation could be that the increased suicidality is a chance finding. However, the use of stratification with the consistent findings within each stratum, as well as the consistency with previous similar findings (Zisook et al., 2004; Zisook et al., 2007a; Zisook et al., 2007b; Thompson 2008) reduce the probability that this was a chance finding.

Strengths and limitations

Our study has several strengths. First, to our knowledge, few previous studies investigated the differences between pre-adult MDD and adult MDD in a naturalistic sample of secondary and tertiary care MDD outpatients. More than 80% of the patients referred for treatment of a MAS disorder between 2004 and 2007 in RD or the LUMC were included in the ROM-database. The external validity of these findings from real-life patients is likely higher than findings in samples from RCTs. Second, the reliability of the results was enhanced by replication of the findings in the two age group samples, ruling out the confounding of age. Most previous studies adjusted for current age (Parker et al., 2003; Ramklint & Ekselius, 2003) or age and duration of illness (e.g. Zisook et al., 2004; Zisook et al., 2007b), which may still leave room for residual confounding. Third, in addition to using a cut-off age of 18 years like in most previous studies, we performed a sensitivity analysis in which we used a cut-off age of 25 years to distinguish between pre-adult and adult onset MDD. This analysis yielded largely comparable results.

A limitation of our study is the fact that the initial onset of MDD was assessed retrospectively, by using self-report estimates, which limits the reliability. This was also the case in many previous studies. This method is more prone to measurement error and bias than assessment through medical records, e.g., a current depressive state can cause recall bias regarding the exact age of onset of the first depressive episode. In addition, the retrospective assessments did not allow for more subtle distinctions, such as the between the onset of first symptoms and the onset of the full syndrome. However, as other studies also assessed age of onset retrospectively, this limitation is not restricted to our study. Other possible limitations include the fact that not all patients were included in the ROM database (e.g. due to language problems) and the fact that excluded patients with missing variables had different mean age, age of onset and MADRS score than included patients. Furthermore, no data regarding familial status, somatic comorbidity, Axis II disorders and information about (pharmacological) treatment was ascertained. Finally, by using the MINI-Plus diagnostic interview we were not able to make a distinction between principal and comorbid diagnoses.

Conclusion

Overall, our findings in a large naturalistic cohort of outpatients with MDD confirm earlier studies that found pre-adult onset MDD associated with more suicidality. Our study stresses the importance of taking the age of onset of MDD in account in clinical decision making and suicide risk assessment.

References

- Aaronson, N.K., Muller, M., Cohen, P.D., Essink-Bot, M.L., Fekkes, M., Sanderman, R., Sprangers, M.A., te Velde, A., Verrips, E., 1998. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. *Journal of Clinical Epidemiology* 51, 1055-1068.
- Alpert, J.E., Fava, M., Uebelacker, L.A., Nierenberg, A.A., Pava, J.A., Worthington, J.J., Rosenbaum, J.F., 1999. Patterns of axis I comorbidity in early-onset versus late-onset major depressive disorder. *Biological Psychiatry* 46, 202-211.
- American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association: Washington, DC.
- Benazzi, F., 2000. Early-onset versus late-onset atypical depression: unipolar and bipolar II. *Journal of Affective Disorders* 61, 95-99.
- de Beurs, E. & Zitman, F.G. 2006. The Brief Symptom Inventory (BSI): Reliability and validity of a practical alternative to SCL-90. *Maandblad Geestelijke Volksgezondheid* 61, 120-141.
- de Beurs, E., den Hollander-Gijsman, M.E., van Rood, Y.R., van der Wee, N.J., Giltay, E.J., van Noorden, M.S., van der Lem, R., van, Fenema. E.M., 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 and Psychotherapy* 18, 1-12.
- Coryell, W., Solomon, D., Leon, A., Fiedorowicz, J.G., Schettler, P., Judd, L., Keller, M., 2009. Does major depressive disorder change with age? *Psychological Medicine* 39, 1689-1695.
- Derogatis, L.R., 1977. Confirmation of the dimensional structure of the SCL-90: A study in construct validation. *Journal of Clinical Psychology* 33, 981-989.
- Derogatis, L.R. & Melisaratos, N., 1983. The Brief Symptom Inventory: an introductory report. *Psychological Medicine* 13, 595-605.
- Fava, M., Alpert, J.E., Borus, J.S., Nierenberg, A.A., Pava, J.A., Rosenbaum, J.F., 1996. Patterns of personality disorder comorbidity in early-onset versus late-onset major depression. *The American Journal of Psychiatry* 153, 1308-1312.
- Forman, E.M., Berk, M.S., Henriques, G.R., Brown, G.K., Beck, A.T., 2004. History of multiple suicide attempts as a behavioural marker of severe psychopathology. *The American Journal of Psychiatry* 161, 437-443.
- Harrington, R., Fudge, H., Rutter, M., Pickles, A., Hill, J., 1990. Adult outcomes of childhood and adolescent depression. I. Psychiatric status. *Archives of General Psychiatry* 47, 465-473.
- Jaffee, S.R., Moffitt, T.E., Caspi, A., Fombonne, E., Poulton, R., Martin, J., 2002. Differences in early childhood risk factors for juvenile-onset and adult-onset depression. *Archives of General Psychiatry* 59, 215-222.
- Kaufman, J., Martin, A., King, R.A., Charney, D., 2001. Are child-, adolescent-, and adult-onset depression one and the same disorder? *Biological Psychiatry* 49, 980-1001.
- Kendler, K.S., Gatz, M., Gardner, C.O., Pedersen, N.L., 2005. Age at onset and familial risk for major depression in a Swedish national twin sample. *Psychological Medicine* 35, 1573-1579.
- Klein, D.N., 1999. Age of onset in chronic major depression: relation to demographic and clinical variables, family history, and treatment response. *Journal of Affective Disorders* 55, 149-157.

- Klein, D.N., Lewinsohn, P.M., Seeley, J.R., Rohde, P., 2001. A family study of major depressive disorder in a community sample of adolescents. *Archives of General Psychiatry* 58, 13-20.
- Kovacs, M., Devlin, B., Pollock, M., Richards, C., Mukerji, P., 1997. A controlled family history study of childhood-onset depressive disorder. *Archives of General Psychiatry* 54, 613-623.
- Lecrubier, Y., Sheehan, D.V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K., Janavs, J., 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.
- Mallinckrodt, C.H., Watkin, J.G., Liu, C., Wohlreich, M.M., Raskin, J., 2005. Duloxetine in the treatment of Major Depressive Disorder: a comparison of efficacy in patients with and without melancholic features. *BMC Psychiatry* 5, 1.
- Malone, K.M., Haas, G.L., Sweeney, J.A., Mann, J.J., 1995. Major depression and the risk of attempted suicide. *Journal of Affective Disorders* 34, 173-185.
- Montgomery, S.A., 1979. A new depression scale designed to be sensitive to change. *The British journal of psychiatry* 134, 382-389.
- Muller, M.J., 2003. Differentiating moderate and severe depression using the Montgomery-Åsberg depression rating scale (MADRS). *Journal of Affective Disorders* 77, 255-260.
- Neuman, R.J., Geller, B., Rice, J.P. & Todd, R.D., 1997. Increased prevalence and earlier onset of mood disorders among relatives of prepubertal versus adult probands. *Journal of the American Academy of Child and Adolescent Psychiatry* 36, 466-473.
- van Noorden, M.S., Giltay, E.J., den Hollander-Gijsman, M.E., van der Wee, N.J., 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.
- Parker, G., Roy, K., Hadzi-Pavlovic, D., Mitchell, P., Wilhelm, K., 2003. Distinguishing early and late onset non-melancholic unipolar depression. *Journal of Affective Disorders* 74, 131-138.
- Ramklint, M. & Ekselius, L., 2003. Personality traits and personality disorders in early onset versus late onset major depression. *Journal of Affective Disorders* 75, 35-42.
- Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV-TR and ICD-10. *Journal of Clinical Psychiatry* 59 Suppl 20, 22-33.
- Snaith, R.P., 1986. Grade scores of the Montgomery-Åsberg Depression and the Clinical Anxiety Scales. *The British Journal of Psychiatry* 148, 599-601.
- Thompson, A.H., 2008. Younger onset of depression is associated with greater suicidal intent. *Social Psychiatry and Psychiatric Epidemiology* 43, 538-544.
- 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.
- van Vliet, I.M., de Beurs, E., 2007. [The MINI-International Neuropsychiatric Interview. A brief structured diagnostic psychiatric interview for DSM-IV-TR en ICD-10 psychiatric disorders]. *Tijdschrift voor Psychiatrie* 49, 393-397.
- Ware, J.E., Sherbourne, C.D., 1992. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care* 30, 473-483.

- Weissman, M.M., Wolk, S., Goldstein, R.B., Moreau, D., Adams, P., Greenwald, S., Klier, C.M., Ryan, N.D., Dahl, R.E., Wickramaratne, P., 1999. Depressed adolescents grown up. *Journal of the American Medical Association* 281, 1707-1713.
- World Health Organisation, 1990. Composite International Diagnostic Interview (CIDI), version 1.1. WHO edn., Geneva.
- Zetin, M. & Hoepner, C.T., 2007. Relevance of exclusion criteria in antidepressant clinical trials: a replication study. *Journal of Clinical Psychopharmacology* 27, 295-301.
- Zimmerman, M., Chelminski, I., Posternak, M.A., 2005. Generalisability of antidepressant efficacy trials: differences between depressed psychiatric outpatients who would or would not qualify for an efficacy trial. *The American Journal of Psychiatry* 162, 1370-1372.
- Zimmerman, M., Mattia, J.I., Posternak, M.A., 2002. Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? *The American Journal of Psychiatry* 159, 469-473.
- Zisook, S., Lesser, I., Stewart, J.W., Wisniewski, S.R., Balasubramani, G.K., Fava, M., Gilmer, W.S., Dresselhaus, T.R., Thase, M.E., Nierenberg, A.A., Trivedi, M.H., Rush, A.J., 2007a. Effect of age at onset on the course of major depressive disorder. *The American Journal of Psychiatry* 164, 1539-1546.
- Zisook, S., Rush, A.J., Albala, A., Alpert, J., Balasubramani, G.K., Fava, M., Husain, M., Sackeim, H., Trivedi, M., Wisniewski, S., 2004. Factors that differentiate early vs. later onset of major depression disorder. *Psychiatry Research* 129, 127-140.
- Zisook, S., Rush, A.J., Lesser, I., Wisniewski, S.R., Trivedi, M., Husain, M.M., Balasubramani, G.K., Alpert, J.E., Fava, M., 2007b. Preadult onset vs. adult onset of major depressive disorder: a replication study. *Acta Psychiatrica Scandinavica* 115, 196-205.

Chapter 4

Prevalence and correlates of lifetime deliberate self-harm and suicidal ideation in naturalistic outpatients: the Leiden Routine Outcome Monitoring Study

Suzanne de Klerk*
Martijn van Noorden*
Anne van Giezen
Philip Spinhoven
Margien den-Hollander-Gijsman
Erik Giltay
Anne Speckens
Frans Zitman

*Joint first authors

Journal of affective disorders (2011), 133, 257-264

Abstract

Background: Deliberate self-harm and suicidal ideation (DSHI) are common phenomena in general and mental health populations. Identifying factors associated with DSHI may contribute to the early identification, prevention and treatment of DSHI. Aims of the study are to determine the prevalence and correlates of lifetime DSHI in a naturalistic sample of psychiatric outpatients with mood, anxiety or somatoform (MAS) disorders.

Methods: Of 3,798 consecutive patients from January 2004 to December 2006, 2,844 (74.9%) patients were analysed (mean age=37.5, SD=12.0; age range: 18-65; 62.7% women). Lifetime DSHI was assessed with Routine Outcome Monitoring (ROM), including demographic parameters, DSM-IV diagnosis, depressive symptoms, symptoms of anxiety, general psychopathology and personality traits.

Results: Of the 2,844 subjects, 55% reported lifetime DSHI. In multivariable logistic regression analysis, the most important factors associated with lifetime DSHI were being unmarried, low education, high number of psychiatric diagnoses, lower anxiety scores, higher depression scores and the personality trait of emotional dysregulation.

Limitations: Deliberate self-harm may have been under-reported in self-report questionnaires; the assessment of personality traits may have been influenced by state psychopathology; traumatic events were not assessed.

Discussion: The findings suggest that DSHI is common among psychiatric outpatients with MAS disorders and that current symptoms and underlying personality vulnerabilities were independently involved in DSHI. Whether symptoms of somatic anxiety are protective should be confirmed in subsequent studies. These findings may help clinicians in identifying patients at risk for deliberate self-harm and suicide.

Introduction

Deliberate self-harm is a common problem in the general population, as well as in psychiatric populations (Briere & Gil, 1998; Klonsky et al., 2003). It has been defined as the intentional self-injury without (conscious) suicidal intent (Chapman et al., 2006; Klonsky et al., 2003), or as behaviour that was self-initiated with the intent to harm the body, regardless of intent to die (De Leo et al., 2004; Schmidtke et al., 2004). In the current study, both non-suicidal and suicidal self-harm, as well as thoughts about self-harm and suicide were analysed (deliberate self-harm or ideation; DSHI).

With regard to deliberate self-harm *behaviour*, a lifetime prevalence of 4% was found in nonclinical populations in the US (Briere & Gil, 1998; Klonsky et al., 2003). Prevalence rates in selected populations were higher, especially for adolescents (7%-9%; Hawton et al., 2002; Skegg, 2005) and psychiatric inpatients (21%-61%; Briere & Gil, 1998; Suyemoto, 1998). In the UK, an estimated 170,000 deliberate self-harm patients per year are admitted at emergency departments of general hospitals with deliberate self-poisoning (Kapur et al., 1998). The most frequent presentations of deliberate self-harm behaviour are skin-cutting and self-poisoning (Hawton et al., 2002; Schnyder et al., 1999). With regard to deliberate self-harm *ideation* in community samples, lifetime prevalence rates of 11%-14% were found for adults (Kessler et al., 1999; Ten Have et al., 2009) and 19% for adolescents (Evans et al., 2005a). Besides the fact that DSHI is a prevalent problem in various populations, persistence is also very common (Owens et al., 2002) and a history of DSHI increases the risk of a completed suicide (Cooper et al., 2005; Hawton et al., 2003; Hawton & Harriss, 2007; Kapur et al., 2006; Suominen et al., 2004; Sinclair et al., 2010).

According to a large adult community survey the factors that were associated with deliberate self-harm and respectively with suicidal ideation were largely identical (Ten Have et al., 2009). Sociodemographic factors such as younger age and female sex were found to be related to DSHI. (Kessler et al., 1999 ; Kessler et al., 2005; Schmidtke et al., 1996; Evans et al., 2005b). Subjects with low socio-economic status (Schmidtke et al., 1996) and divorced subjects (Kessler et al., 1999; Petronis et al., 1990) were more likely to engage in DSHI.

Disease-related predictors of DSHI in community and clinical populations were the presence of psychiatric symptoms and disorders, such as major depression, substance abuse and anxiety disorders (Kessler et al., 1999; Haw et al., 2001; Plener et al., 2009). With regard to psychiatric symptoms, adolescents with DSHI showed higher levels of depressive (Hawton et al., 2002) and anxiety symptoms, and used more alcohol (Haavisto et al., 2005; Ross and Heath, 2002; Tuisku et al., 2009).

General vulnerability factors such as personality traits and emotion regulation difficulties were found to be related to DSHI in community and clinical samples. The presence of DSHI was increased in patients with personality disorders, especially in patients with borderline personality disorder (Shearer et al., 1988; Haw et al., 2001). Some of the Big-Five personality traits (neuroticism, agreeableness, extraversion, conscientiousness and openness to experience) were also described to be related to lifetime DSHI. In several community sample studies neuroticism was significantly higher in patients with a history of DSHI compared to non-DSHI patients (Brown et al., 2007; Kessler et al., 1999; Klonsky et al., 2003; Williams & Hassanyeh, 1983). Brown found higher levels of neuroticism and openness to experience, and lower levels of agreeableness and conscientiousness in college students who had a history of (non-suicidal) self-harm compared to controls (Brown et al., 2009). Additionally, in a study by Klonsky et al., military recruits reporting a history of self-harm showed more personality pathology traits (borderline, schizotypal, dependent and avoidant) compared to the non-self-harm group (Klonsky et al., 2003). They also found a coherent 'self-harm personality profile', according to self- and peer reports: individuals with a history of self-harm tended to have strange and intense emotions and a heightened sensitivity to interpersonal rejection. 'Emotion dysregulation' difficulties in responding to one's own emotions as a personality trait, may play a central role in DSHI (Chapman et al., 2006; Linehan, 1993; Gratz & Roemer, 2008). A study of DSHI risk factors in male students (Gratz & Chapman, 2007) found emotion dysregulation to distinguish between men with and without DSHI.

Most studies of DSHI focus on the general community, students or inpatients. Knowledge about DSHI-related factors regarding outpatients with common mental disorders such as mood, anxiety and somatoform disorders in a naturalistic mental health care setting is scarce. Identifying associative factors of DSHI is of clinical importance and may assist clinicians, social workers, prevention workers and policy makers in the identification, prevention and treatment of DSHI and eventually, in preventing suicide. In this explorative study, we aimed to investigate the prevalence rate of lifetime DSHI and its demographic and disease specific correlates in a large Dutch naturalistic outpatient mental health population.

Methods

Subjects and procedure

The sample consisted of 3,798 outpatients from the Dutch Regional Mental Health Provider (RMHP) Rivierduinen and the psychiatric outpatient department of the Leiden University Medical Center (LUMC). Patients were referred to one of these locations for treatment of a mood, anxiety or somatoform (MAS) disorder between January 2004 and December 2006. All outpatients that enter the clinics were routinely assessed with Routine Outcome Monitoring (ROM), as part of the usual diagnostic procedure (De Beurs et al., 2010). ROM consists of an extensive battery of psychometric instruments, both self-report and interviewer-based. All interviewer-based measurements were carried out by trained research nurses or psychologists. Self-report questionnaires were completed using a touch-screen computer. The assessment took about 120 minutes. Only patients with insufficient mastery of the Dutch language and patients who were unable to complete computerised and written questionnaires were ineligible for ROM. Approximately, 80% of the referred MAS patients were assessed with ROM in the study period. For the current study, we used baseline ROM assessments only.

Of the 3,798 outpatients, 2,844 patients between the age of 18 to 65 had complete data on variables of interest and were analysed (74.9%; Figure 4.1). In total, 116 patients were excluded because of their age and 838 patients were excluded because they had missing data. The main reason for missing data ($N=395$) was that the Dimensional Assessment of Personality Pathology-Basic Questionnaire Short Form (DAPP-SF) had been introduced in ROM gradually. Mean age of the excluded patients was significantly higher in comparison with the included group ($M=39.6$ vs. 37.5 , $p<0.001$). The excluded patients had significantly higher scores on general psychopathology (BSI; $M=1.4$ vs. 1.2 , $p<0.001$), depression and anxiety (MADRS; $M=19.9$ vs. 17.7 , $p<0.001$, BAS; $M=15.8$ vs. 14.2 , $p<0.001$). The groups also differed on nationality and employment status (in the excluded patients less people were Dutch, $p<0.001$ or employed, $p<0.001$). Finally, excluded patients were more likely to be married or had a partner than the included patients ($p<0.05$). Concerning DSHI ($N=448$), education, number of diagnoses, gender, housing and suicide attempt history no differences were found between the included and excluded patients. Of the final 2,844 patients, 1,061 patients were male (37.3 %; mean age= 38.8 , $SD=11.8$) and 1,783 were female (62.7 %; mean age= 36.7 , $SD=12.0$).

Measures

Demographics: For the current study, demographic variables were obtained using a self-report questionnaire that assessed age, gender, ethnic background, education, marital status, housing situation and employment status. A Dutch ethnic background was assumed when the patient and both parents were born in the Netherlands. 'Other ethnicity' was scored when these criteria were not fulfilled.

DSHI and personality pathology: To assess lifetime DSHI and personality pathology, the DAPP-SF was used (Van Kampen et al., 2008). The DAPP-SF contains 136 items and consists of 18 subscales covering clinically relevant dimensions of personality pathology, namely: Submissiveness, Cognitive Distortion, Identity Problems, Affective Lability, Stimulus Seeking, Compulsivity, Restricted Expression, Callousness, Oppositionality, Intimacy Problems, Rejection, Anxiousness, Conduct Problems, Suspiciousness, Social Avoidance, Narcissism, Insecure Attachment and Self-Harm. The latter subscale was used to measure DSHI.

There are four higher-order constructs containing several subscales, namely Emotional Dysregulation, Dissocial Behaviour, Inhibition and Compulsivity. The items of the DAPP-SF are rated on a 5-point Likert scale, with scores ranging from 1 ("very unlike me") to 5 ("very like me"). A high total (and subscale) score indicates the presence of (a) deviant personality trait(s). The Self-Harm subscale consists of 6 items and measures lifetime self-harm ideation and behaviour, specifically a history of self-poisoning (1 item), repetitive self-injury (1 item) and suicidal ideation and intent (4 items) ranging from 1 to 5 (mean score). The DAPP-SF has good psychometric properties. Cronbach's alpha of the subscale Self-Harm in a previous sample with MAS disorders was 0.88. and respectively 0.90 in a sample with personality disorders. Furthermore, the subscale correlated strong with general psychopathology (De Beurs et al., 2009). Cronbach's alpha in the current sample was respectively 0.96 for the total DAPP-SF, 0.89 for the DSHI subscale, 0.97 for emotional dysregulation, 0.88 for dissocial behaviour, 0.80 for inhibition and 0.84 for compulsivity. Also the subscales had satisfactory internal consistencies, ranging from 0.77 (conduct problems) to 0.91 (suspiciousness). We defined the presence of self-harm as a score of ≥ 2 on the self-injury and/or the self-poisoning item. The presence of self-harm ideation was defined if subject scored ≥ 2 on one or more suicidal ideation and intent items.

Psychiatric diagnosis: Diagnostic status according to the Diagnostic and Statistical Manual of Mental Disorders IV-TR (American Psychiatric Association, 2000) was assessed with the Dutch version of the MINI International Neuropsychiatric Interview Plus, version 5.00-R; MINI-Plus, a standardised diagnostic interview (Sheehan et al., 1998; Van

Vliet & De Beurs, 2007). The interview consists of 23 modules in which the presence or absence of DSM-IV criteria for the main psychiatric disorders was examined. The presence of a 'prior suicide attempt' and the amount of Axis I disorders was also assessed with the MINI-Plus. The MINI-plus was tested in different samples and appeared to have good psychometric properties (Sheehan et al., 1998; Lecrubier et al., 1997).

Psychiatric symptoms: Depressive and anxiety symptoms were measured by the Montgomery-Åsberg Depression Rating Scale (MADRS; Montgomery & Åsberg, 1979) and the Brief Anxiety Scale (BAS; Tyrer et al., 1984). The MADRS and the BAS have good psychometric properties (Goekoop et al., 1994). Cronbach's alpha for the MADRS in the current sample was 0.86 and for the BAS 0.68. General Psychopathology was measured by the Brief Symptom Inventory (BSI; Derogatis, 1975). This is a shortened version of the Symptom Checklist (SCL-90) and contains 53 self-report items. The BSI measures current symptoms of psychopathology in adults covering nine symptom dimensions and three global indices of distress (Derogatis, 1975; De Beurs & Zitman, 2006). The respondent indicates to what extent they have been bothered by a symptom in the past week, including today, with scores ranging from 0 ("not at all") to 5 ("extremely"). The Dutch translation of the BSI has good psychometric properties (De Beurs & Zitman, 2006). In the current sample, the Cronbach's alpha for the BSI was 0.97.

Statistical Analyses

Patients were categorised in three groups with no, intermediate and high DSHI (i.e., tertiles) in order to increase clarity and interpretability of the results. The chi-squared test for linear-by-linear association or Kruskal-Wallis nonparametric test were used to analyse associations with categorical and ordinal variables, respectively. A prediction model was built in which contributing factors to DSHI were added to a multinomial logistic regression analysis, and the potential predictive factors were chosen by univariable analysis ($p < 0.1$), while age and gender were forced into the model. In Table 4.4, associations are presented as odds ratios (OR) with 95% confidence intervals (95% CI) per one standard deviation increase in predictive factor (per 1-SD using z-scores; to facilitate the comparison between different predictive factors), in which no DSHI was considered the reference category. The associations with predictor factors were examined using multivariable linear regression analysis, and these associations are presented as the p-value for linear trend over the tertiles of DSHI (i.e., 1, 2 or 3). In addition, forward stepwise linear regression analysis was used to find the minimum subset of independent predictive factors of DSHI from the longer list of possible factors. In this final model, the stringent significance levels of 0.01 for removal and of 0.005 for entry were chosen, and age and sex were forced into

the model. A two sided p-value of <0.05 was considered statistically significant, and the statistical analyses were performed using SPSS version 17.0 statistical software.

Results

Descriptives

In total, 1,549 patients (54.5%; 95% confidence interval [CI]: 52.6%-56.3%) reported lifetime DSHI, based on the tertile distribution of the DAPP-SF. The sample consisted of 1,295 (45.5%) patients with no DSHI, 597 patients (21.0%) with intermediate DSHI and 952 patients (33.5%) with high DSHI. Because of the large number of subjects with no DSHI, the tertiles were not equally distributed. Of the total 2,844 patients, 689 (24.2%) reported a history of self-harm or suicide attempt whereas 1,433 (50.4%) reported a history of suicidal ideation. 573 subjects (20.1%) reported suicidal ideation *and* self-harm or a suicide attempt. According to the MINI International Neuropsychiatric Interview-Plus (MINI-Plus) 411 subjects (14.5%) reported a prior suicide attempt. Of the 2,844 patients 1,300 (45.7%) had a current depression or dysthymia, 1,224 (43.0%) a current anxiety disorder, 368 (12.9%) a current somatoform disorder, and 672 (23.6%) had no MAS disorder. Comorbidity of MAS disorders occurred in 748 (26.3%) patients.

Demographic variables and DSHI

Table 4.1 presents demographic characteristics of the sample according to DSHI group. Linear-by-linear association revealed no significant trend for gender and DSHI group. Linear-by-linear association tests revealed a significant increasing association of other ethnicity from no DSHI (14.2%), to intermediate DSHI (19.4%) and high DSHI (19.7%), $\chi^2(1, 2844)=12.60$, $p<0.001$, and respectively for being unmarried with DSHI group, $\chi^2(1, 2844)=65.65$, $p<0.001$, living alone or with family with DSHI group, $\chi^2(2, 2844)=55.77$, $p<0.001$, lower education with DSHI group, $\chi^2(1, 2844)=6.67$, $p<0.05$, being unemployed with DSHI group, $\chi^2(1, =2844)=41.22$, $p<0.001$ and younger age with DSHI group, $F(1, 2841)=8.62$, $p<0.005$.

Table 4.1. Sociodemographic characteristics in total sample and according to level of DSHI in 2,844 outpatients

	Total sample	Tertiles of DSHI			p-value (linear-by-linear)
	(n=2844, 100%)	No DSHI (n=1295, 45.5%)	Intermediate DSHI (n=597, 21.0%)	High DSHI (n=952, 33.5%)	
DSHI median (range)	1.2 (1-5)	1 (1-1)	1.3 (1.2-1.7)	2.7 (1.8-5)	-
Age (mean $\pm$ SD)	37.5 $\pm$ 12.0	38.2 $\pm$ 12.0	37.4 $\pm$ 11.6	36.7 $\pm$ 12.0	0.003
Gender (n, %)					0.510
Male	1061 (37.3)	466 (36.0)	242 (40.5)	353 (37.1)	
Female	1783 (62.7)	829 (64.0)	355 (59.5)	599 (62.9)	
Ethnic Background (n, %)					<0.001
Dutch	2356 (82.8)	1111 (85.8)	481 (80.6)	764 (80.3)	
Other ethnicity	488 (17.2)	184 (14.2)	116 (19.4)	188 (19.7)	
Marital status (n, %)					<0.001
Married	1432 (50.4)	743 (57.4)	309 (51.8)	380 (39.9)	
Unmarried	1412 (49.6)	552 (42.6)	288 (48.2)	572 (60.1)	
Housing situation (n, %)					<0.001
Living with partner	1465 (51.5)	747 (57.7)	323 (54.1)	395 (41.5)	
Living alone or with family	1379 (48.5)	548 (42.3)	274 (45.9)	557 (58.5)	
Educational status (n, %)					0.01
Higher education	1219 (42.9)	757 (58.5)	367 (61.5)	501 (52.6)	
Lower education	1625 (57.1)	538 (41.5)	230 (38.5)	451 (47.4)	
Employment status (n, %)					<0.001
Employed	1331 (46.8)	680 (52.5)	282 (47.2)	369 (38.8)	
Unemployed	1513 (53.2)	615 (47.5)	315 (52.8)	583 (61.2)	

Abbreviations: DSHI: Deliberate Self-Harm or Ideation; SD: Standard Deviation

Disease-related variables: psychiatric disorders and DSHI

Table 4.2 shows psychiatric disorders and symptoms according to the level of DSHI. Significant linear-by-linear associations were found for all MAS disorders with DSHI group except for dysthymia. Linear-by-linear associations revealed a significant association between increasing prevalence rates of depression and anxiety disorders for each subsequent DSHI group, except for panic disorder, generalised anxiety disorder and somatoform disorder, where the test revealed a significant inverse association. Additionally, a significant increasing association was found for alcohol and drug abuse or dependence from no DSHI (2.6%; 2.2% for alcohol and drug abuse or dependence, respectively), intermediate DSHI (4.7%; 3.9%) to high DSHI (8.7%; 7.4%), $\chi^2(1, 2844)=41.41$, $p<0.001$; $\chi^2(1, 2844) =35.63$, $p<0.001$). Linear-by-linear associations also revealed a significant increasing association of the number of diagnoses, $F(1, 2841)=111.33$, $p<0.001$, depressive symptoms (MADRS), $F(1, 2841)=484.14$, $p<0.001$, symptoms of anxiety (BAS), $F(1, 2841)=76.37$, $p<0.001$, and general psychopathology, $F(1, 2841)=569.42$, $p<0.001$ from no DSHI, to intermediate and high DSHI.

Table 4.2. Psychiatric disorders (MINI-diagnoses) and symptoms in total sample and according to the level of DSHI in 2844 outpatients

	Total sample (n=2844, 100%)	Teriles of DSHI			p-value for trend (linear-by-linear) ^a
		No DSHI (n=1295, 45.5%)	Intermediate DSHI (n=597, 21.0%)	High DSHI (n=952, 33.5%)	
Psychiatric disorders (n, %)					
Mood disorder					
Depression	1191 (41.9)	364 (28.1)	241 (40.4)	586 (61.6)	<0.001
Dysthymic disorder	109 (3.8)	43 (3.3)	26 (4.4)	40 (4.2)	0.26
Anxiety disorder					
Posttraumatic stress disorder	286 (10.1)	96 (7.4)	54 (9.0)	136 (14.3)	<0.001
Social phobia	353 (12.4)	130 (10.0)	83 (13.9)	141 (14.7)	0.001
Obsessive compulsive disorder	133 (4.7)	44 (3.4)	36 (6.0)	53 (5.6)	0.012
Panic disorder	334 (11.7)	177 (13.7)	64 (10.7)	93 (9.8)	0.004
Generalized anxiety disorder	163 (5.7)	82 (6.3)	41 (6.9)	40 (4.2)	0.041
Any somatoform disorder	461 (16.2)	255 (19.7)	80 (13.4)	126 (13.2)	<0.001
Alcohol abuse or dependence	145 (5.1)	34 (2.6)	28 (4.7)	83 (8.7)	<0.001
Drug abuse or dependence	121 (4.3)	28 (2.2)	23 (3.9)	70 (7.4)	<0.001
Prior suicide attempt (MINI)	411 (14.5)	39 (3.0)	80 (13.4)	292 (30.7)	<0.001
Total number of diagnoses (mean ± SD)	2.0 ± 1.2	1.7 ± 1.1	2.0 ± 1.2	2.3 ± 1.3	<0.001
Psychiatric symptoms (mean ± SD)					
Depressive symptoms (MADRS)	17.7 ± 9.8	14.1 ± 8.6	17.7 ± 9.2	22.7 ± 9.5	<0.001
Symptoms of anxiety (BAS)	14.2 ± 6.8	13.1 ± 6.6	14.6 ± 7.0	15.6 ± 6.7	<0.001
General Psychopathology (BSI)	1.2 ± 0.7	0.9 ± 0.6	1.2 ± 0.7	1.5 ± 0.7	<0.001

Abbreviations: DSHI: Deliberate Self-Harm or Ideation; MADRS: Montgomery-Åsberg Depression Rating Scale; BAS: Brief Anxiety Scale; BSI: Brief Symptom Inventory.

^a p-value for trend by Kruskal Wallis or chi-squared test.

Table 4.3. Personality characteristics in total sample and according to the level of DSHI in 2,844 outpatients

	Total sample	Tertiles of DSHI			p-value for trend ^a
	(n=2844, 100%)	No DSHI (n=1295, 45.5%)	Intermediate DSHI (n=597, 21.0%)	High DSHI (n=952, 33.5%)	
Total DAPP-SF (mean ± SD)	2.5 ± 0.5	2.3 ± 0.5	2.6 ± 0.5	2.9 ± 0.5	<0.001
Personality Constructs (mean ± SD)					
Emotional dysregulation	2.7 ± 0.7	2.4 ± 0.7	2.8 ± 0.6	3.1 ± 0.6	<0.001
Dissocial behavior	1.9 ± 0.5	1.8 ± 0.5	1.9 ± 0.5	2.1 ± 0.6	<0.001
Inhibitedness	2.7 ± 0.7	2.6 ± 0.6	2.8 ± 0.6	3.0 ± 0.7	<0.001
Compulsivity	2.8 ± 0.9	2.7 ± 0.9	2.9 ± 0.9	2.9 ± 0.9	<0.001

Abbreviations: DSHI: Deliberate Self-Harm or Ideation; DAPP-SF: Dimensional Assessment of Personality Pathology -short form; Total DAPP-SF excludes the self-harm scale items.

^a p-value for trend by Kruskal Wallis .

Personality characteristics and DSHI

Table 4.3 shows personality characteristics according to the level of DSHI. Linear-by-linear associations revealed a significant increasing association of total DAPP-SF score from no DSHI, to intermediate and high DSHI, $F(1, 2841) = 788.20, p < 0.001$. Significant increasing associations were also found for the personality constructs emotion dysregulation, $F(1, 2841) = 786.99, p < 0.001$, dissocial behaviour, $F(1, 2841) = 153.14, p < 0.001$, inhibitedness, $F(1, 2841) = 164.68, p < 0.001$ and compulsivity, $F(1, 2841) = 27.12, p < 0.001$.

Multivariable analysis

Table 4.4 presents the results of a multinomial logistic regression analysis. The overall model significantly predicted DSHI, $R^2 = 0.35$ (Nagelkerke), $\chi^2(28, 28) = 1056, p < 0.001$. A forward stepwise linear regression analyses revealed significant positive trends of DSHI for low education, $\beta = 0.08$; unmarried or no partner, $\beta = 0.05$; number of diagnoses, $\beta = 0.07$; symptoms of anxiety, $\beta = -0.27$; depressive symptoms, $\beta = 0.39$; and emotion dysregulation, $\beta = 0.36$.

Discussion

According to the findings of the present study, deliberate self harm and suicidal ideation are prevalent problems in psychiatric outpatients with MAS disorders, with 55% of the patients reporting thoughts or behaviours concerning deliberately injuring one’s own body. This prevalence rate is comparable to previous studies among *depressed adolescent* outpatients where prevalence rates between 42% and 58% were found for any suicidality, including thoughts and behaviour (Tuisku et al., 2006; Pelkonen et al., 1997). This study

Table 4.4 Predictors of DSH-I in multinomial logistic regression analysis in 2,844 psychiatric outpatients

	Tertiles of DSH					
	No DSH-I (ref) (N=597)	Intermediate DSH-I (OR, 95% CI) (N=952)	High DSH-I (OR, 95% CI) (N=1295)	t	Multivariable p-value for trend ^a	t
Age	1.0	1.00 (0.99-1.01)	1.00 (0.99-1.01)	-0.33	0.74	-0.58
Male sex	1.0	1.28 (1.02-1.61)	1.02 (0.81-1.28)	-0.70	0.48	-1.90
Dutch background	1.0	0.80 (0.61-1.05)	0.94 (0.72-1.24)	0.35	0.72	-
Married or partner	1.0	0.93 (0.75-1.15)	0.65 (0.52-0.80)	3.90	<0.001	4.69
High education	1.0	1.09 (0.88-1.35)	0.73 (0.59-0.90)	2.83	0.005	3.39
Employed	1.0	0.87 (0.70-1.07)	0.73 (0.62-0.94)	2.50	0.012	-
Number of diagnoses (MIN)	1.0	1.05 (0.94-0.17)	1.22 (1.09-1.36)	3.47	0.001	3.87
Symptoms of anxiety (BAS)	1.0	0.76 (0.65-0.90)	0.40 (0.34-0.47)	-11.75	<0.001	-11.87
General psychopathology (BSI)	1.0	1.20 (0.99-1.46)	1.20 (1.00-1.45)	2.25	0.025	-
Depressive symptoms (MADRS)	1.0	1.35 (1.13-1.61)	3.13 (2.63-3.73)	13.73	<0.001	16.72
Emotion dysregulation (DAPP)	1.0	1.66 (1.41-1.96)	2.69 (2.27-3.19)	12.11	<0.001	19.58
Dissocial behavior (DAPP)	1.0	1.07 (0.94-1.22)	1.22 (1.08-1.38)	3.04	0.002	-
Inhibition (DAPP)	1.0	1.08 (0.96-1.21)	1.12 (1.00-1.25)	1.84	0.07	-
Compulsivity (DAPP)	1.0	0.96 (0.86-1.08)	0.95 (0.85-1.05)	-1.33	0.18	-

Abbreviations: DSH-I: Deliberate self-harm or Ideation; DAPP: Dimensional Assessment of Personality Pathology-Short Form; MADRS: Montgomery-Åsberg Depression Rating Scale; BAS: Brief Anxiety Scale; BSI: Brief Symptom Inventory; MINI: Mini International Neuropsychiatric Interview Plus; CI: Confidence Interval; OR: Odds Ratio; ORs are presented as z-scores.

^a forced entry in multivariable linear regression;

^b forward stepwise linear regression analysis; $p=0.01$ for removal and $p=0.005$ for entry.

confirmed the hypotheses that demographic factors, disease-related factors and general vulnerability factors are associated to DSHI in psychiatric outpatients with MAS disorders. We found that being unmarried, lower education, number of psychiatric diagnoses, lower anxiety scores, higher depression scores and emotional dysregulation were strong independent predictors of DSHI.

We found several factors associated to lifetime DSHI. First, sociodemographic factors such as being unmarried, unemployed and lower education were significantly related to the presence of DSHI in univariable analyses. In multivariable analyses, being unmarried and lower education were independent predictors of DSHI. These findings are comparable to previous research where social disadvantages such as no work and lower education are seen as risk factors for mental disorders and suicide attempts (Kapur et al., 2006; Taylor et al., 2004). A finding that contrasts with some earlier research is the lack of significant gender differences in the prevalence of DSHI (Herpertz, 1995). These results contribute to the discussion whether DSHI may be as common among men as it is among women (Muehlenkamp & Gutierrez, 2004).

Secondly, disease-related factors, specifically depressive and anxiety symptoms, general psychopathology, and number of diagnoses were positively associated with lifetime DSHI in univariable analyses. Consistent with previous research (Fliege et al., 2006) DSHI patients reported a higher degree of symptoms of depression, anxiety and general psychopathology (measured with the MADRS, BAS and BSI, respectively). In multivariable analyses, higher depression scores and number of diagnoses predicted DSHI. Surprisingly, when controlled for the other variables in the multivariable analysis, anxiety symptoms measured with the BAS showed an inverse association with DSHI. In previous research anxiety scales in the social phobia domain were related to lower levels of suicidal behaviour after controlling for depression and suicidal ideation (Valentiner et al., 2002). In another study nonattempters had significantly higher scores on anxiety than attempters of suicide (Placidi et al., 2000). In our sample, social phobia, posttraumatic stress disorder and obsessive compulsive disorder were more prevalent in the high DSHI group in univariable analyses, whereas panic disorder and generalised anxiety disorder were less prevalent in the high DSHI group. In this study, anxiety was measured with the BAS, a scale assessing mainly the somatic aspects of anxiety (Keedwell & Snaith, 1996). When corrected for depression in multivariable analyses, these somatic aspects of anxiety could be protective for DSHI. This is consistent with the finding that in our univariable analyses, the presence of a somatoform disorder showed a decreasing trend over the DSHI groups. Future research might examine whether this results can be replicated in other samples, and with other measures of anxiety.

Thirdly, with regard to general vulnerability factors, especially emotional dysregulation was strongly associated with lifetime DSHI, in both univariable and multivariable analyses. Several researchers claim that DSHI may in fact be an emotion regulation strategy itself, suggesting that DSHI patients have difficulties in coping with distress (Nock & Prinstein, 2004; Slee et al., 2008b) and that DSHI can be a way to avoid emotional problems (Slee et al., 2008a, b). Thereby, Brown et al. (2007) found that college-students with a past history of DSHI and with a recent history of DSHI both reported more negative emotions than the non-DSHI patients. This may indicate a general emotional vulnerability in patients with a history of DSHI.

Strengths and limitations

Previous studies on prevalence and correlates of DSHI often focused on specific disorders (personality, anxiety or depressive disorders), or specific populations such as adolescents or inpatients. Strengths of this study are the large naturalistic psychiatric outpatient study population of 2,844 patients, the routinely adapted assessments by specifically trained research nurses and psychologists and the wide range of potential contributing factors assessed in the study, which focuses on current symptoms and disorders as well as on underlying vulnerabilities such as personality traits.

The study also has several potential limitations. First, not all referred MAS patients were assessed, because about 20% of all patients did not have a baseline ROM assessment due to language problems or logistical reasons. A second limitation includes the assessments of DSHI and personality traits. The DAPP-SF self-harm subscale measures lifetime suicidal ideation (4 items) and specifically self-poisoning and repetitive self-injury. Only two items specifically assess lifetime self-harming behaviour. According to previous research, specific deliberate self-harm behaviours (such as cutting, scratching or head banging) may be underreported according to self-report questionnaires unless they are specifically asked for (Gratz, 2001). Additionally personality traits were assessed by a self-report questionnaire. This may lead to bias such as social desirability and the predisposition to self-enhancement (Herpertz, 1995; Kocalevent et al., 2005; Paulhus & Vazire, 2007). Furthermore, the assessment of personality traits may be influenced by current symptoms of anxiety, depression and general psychopathology. Finally, in the current study, traumatic events were not assessed. Trauma was associated with deliberate self-harm behaviour and ideation in previous studies (Gratz et al., 2002; Romans et al., 1995).

In conclusion, with these results from psychiatric outpatients with MAS disorders about the factors associated with DSHI it may be easier to identify patients at risk for deliberate self-harm and suicide. The overall findings may implicate that interventions

for DSHI should not primarily focus on associated mental disorders such as depression, but should also focus on DSHI itself and specifically target emotion regulation strategies. Whether symptoms of somatic anxiety are protective should be confirmed in subsequent studies. Finally, the findings suggest that both present symptoms as well as underlying personality vulnerabilities seem to be involved in DSHI. Treatment should not merely focus on either personality pathology or for example depressive symptoms: an integrated approach seems more promising (Slee et al., 2008a, b; Slee et al., 2007; Weinberg et al., 2010). Specific therapeutic implications of these findings should be further studied.

References

- American Psychiatric Association, 2000. Diagnostic and statistical manual of mental disorders. American Psychiatric Association: Washington, DC.
- de Beurs, E., Zitman, F. G., 2006. The Brief Symptom Inventory (BSI): Reliability and validity of a practical alternative to SCL-90. *Maandblad Geestelijke Volksgezondheid* 61, 120-141.
- de Beurs, E., Rinne, T., van Kampen, D., Verheul, R., Andrea, H., 2009. Reliability and validity of the Dutch Dimensional Assessment of Personality Pathology-Short Form (DAPP-SF), a shortened version of the DAPP-Basic Questionnaire. *Journal of Personality Disorders* 23, 308-326.
- de Beurs E., den Hollander-Gijsman, M.E., van Rood, Y.R., van der Wee, N.J., Giltay, E.J., van Noorden, M.S., van der Lem, R., van Fenema, E.M., 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 and Psychotherapy* 18, 1-12.
- Briere, J., Gil, E., 1998. Self-mutilation in clinical and general population samples: prevalence, correlates, and functions. *American Journal of Orthopsychiatry* 68, 609-620.
- Brown, S.A., 2009. Personality and non-suicidal deliberate self-harm: Trait differences among a non-clinical population. *Psychiatry Research* 169, 28-32.
- Brown, S.A., Williams, K., Collins, A., 2007. Past and recent deliberate self-harm: emotion and coping strategy differences. *Journal of Clinical Psychology* 63, 791-803.
- Chapman, A.L., Gratz, K.L., Brown, M.Z., 2006. Solving the puzzle of deliberate self-harm: the experiential avoidance model. *Behaviour Research and Therapy* 44, 371-394.
- Cooper, J., Kapur, N., Webb, R., Lawlor, M., Guthrie, E., Kway-Jones, K., Appleby, L., 2005. Suicide after deliberate self-harm: a 4-year cohort study. *The American Journal of Psychiatry* 162, 297-303.
- DeLeo, D., Bille-Brahe, U., Kerkhof, A., Schmidtke, A., 2004. Suicidal behaviour. Hogrefe & Huber, Gottingen.
- Derogatis, L. R., 1975. The Brief Symptom Inventory. Clinical Psychometric Research, Baltimore, MD.
- Evans, E., Hawton, K., Rodham, K., Deeks, J., 2005a. The prevalence of suicidal phenomena in adolescents: a systematic review of population-based studies. *Suicide and Life Threatening Behaviour* 35, 239-250.
- Evans, E., Hawton, K., Rodham, K., 2005b. In what ways are adolescents who engage in self-harm or experience thoughts of self-harm different in terms of help-seeking, communication and coping strategies? *Journal of Adolescence* 28, 573-587.
- Fliege, H., Kocalevent, R.D., Walter, O.B., Beck, S., Gratz, K.L., Gutierrez, P.M., Klapp, B.F., 2006. Three assessment tools for deliberate self-harm and suicide behaviour: evaluation and psychopathological correlates. *Journal of Psychosomatic Research* 61, 113-121.
- Goekoop, J.G., Knoppert-van der Klein, E.A.M., Hoeksema, T., Zwinderman, A.H., 1994. The Dutch version of the Comprehensive Psychopathological Rating Scale. Reliability, factor structure and intensity rating scales. *Tijdschrift voor Psychiatrie* 36, 520-527.
- Gratz, K.L., 2001. Measurement of deliberate self-harm: preliminary data on the deliberate self-harm inventory. *Journal of Psychopathology and Behavioural Assessment* 23, 253-263.
- Gratz, K.L., Chapman, A.L., 2007. The role of emotional responding and childhood maltreatment in the development and maintenance of deliberate self-harm among male undergraduates. *Psychology of Men and Masculinity* 8, 1-14.

- Gratz, K.L., Conrad, S.D., Roemer, L., 2002. Risk factors for deliberate self-harm among college students. *American Journal of Orthopsychiatry* 72, 128-140.
- Gratz, K.L., Roemer, L., 2008. The relationship between emotion dysregulation and deliberate self-harm among female undergraduate students at an urban commuter university. *Cognitive Behaviour Therapy* 37, 14-25.
- Haavisto, A., Sourander, A., Multimaki, P., Parkkola, K., Santalahti, P., Helenius, H., Nikolakaros, G., Moilanen, I., Kumpulainen, K., Piha, J., Aronen, E., Puura, K., Linna, S.L., Almqvist, F., 2005. Factors associated with ideation and acts of deliberate self-harm among 18-year-old boys. A prospective 10-year follow-up study. *Social Psychiatry and Psychiatric Epidemiology* 40, 912-921.
- Haw, C., Hawton, K., Houston, K., Townsend, E., 2001. Psychiatric and personality disorders in deliberate self-harm patients. *The British Journal of Psychiatry* 178, 48-54.
- Hawton, K., Harriss, L., 2007. Deliberate self-harm in young people: characteristics and subsequent mortality in a 20-year cohort of patients presenting to hospital. *Journal of Clinical Psychiatry* 68, 1574-1583.
- Hawton, K., Rodham, K., Evans, E., Weatherall, R., 2002. Deliberate self-harm in adolescents: self-report survey in schools in England. *British Medical Journal* 325, 1207-1211.
- Hawton, K., Zahl, D., Weatherall, R., 2003. Suicide following deliberate self-harm: long-term follow-up of patients who presented to a general hospital. *The British Journal of Psychiatry* 182, 537-542.
- Herpertz, S., 1995. Self-injurious behaviour. Psychopathological and nosological characteristics in subtypes of self-injurers. *Acta Psychiatrica Scandinavica* 91, 57-68.
- 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.
- Kapur, N., Cooper, J., King-Hele, S., Webb, R., Lawlor, M., Rodway, C., Appleby, L., 2006. The repetition of suicidal behaviour: a multicenter cohort study. *Journal of Clinical Psychiatry* 67, 1599-1609.
- Kapur, N., House, A., Creed, F., Feldman, E., Friedman, T., Guthrie, E., 1998. Management of deliberate self-poisoning in adults in four teaching hospitals: descriptive study. *British Medical Journal* 316, 831-832.
- Keedwell, P., Snaith, R.P., 1996. What do anxiety scales measure? *Acta Psychiatrica Scandinavica* 93, 177-180.
- Kessler, R.C., Berglund, P., Borges, G., Nock, M., Wang, P. S., 2005. Trends in suicide ideation, plans, gestures, and attempts in the United States, 1990-1992 to 2001-2003. *Journal of the American Medical Association* 293, 2487-2495.
- Kessler, R.C., Borges, G., Walters, E.E., 1999. Prevalence of and risk factors for lifetime suicide attempts in the National Comorbidity Survey. *Archives of General Psychiatry* 56, 617-626.
- Klonsky, E.D., Oltmanns, T.F., Turkheimer, E., 2003. Deliberate self-harm in a nonclinical population: prevalence and psychological correlates. *American Journal of Psychiatry* 160, 1501-1508.
- Kocalevent, R.D., Fliege, H., Rose, M., Walter, M., Danzer, G., Klapp, B.F., 2005. Autodestructive syndromes. *Psychotherapy and Psychosomatics* 74, 202-211.
- Lecrubier, Y., Sheehan, D.V., Weiller, E., Amorim, P., Bonora, I., Harnett Sheehan, K., Janavs, J., 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.

- Linehan, M. M., 1993. Cognitive-behavioural treatment of borderline personality disorder. Guilford Press, New York, NY.
- Montgomery, S.A., Åsberg, M., 1979. A new depression scale designed to be sensitive to change. *The British Journal of Psychiatry* 134, 382-389.
- Muehlenkamp, J.J., Gutierrez, P.M., 2004. An investigation of differences between self-injurious behaviour and suicide attempts in a sample of adolescents. *Suicide and Life Threatening Behaviour* 34, 12-23.
- Nock, M.K., Prinstein, M.J., 2004. A functional approach to the assessment of self-mutilative behaviour. *Journal of Consulting and Clinical Psychology* 72, 885-890.
- Owens, D., Horrocks, J., House, A., 2002. Fatal and non-fatal repetition of self-harm. Systematic review. *British Journal of Psychiatry* 181, 193-199.
- Paulhus, D. L., Vazire, S., 2007. The self-report method. In: Robins, R. W., Fraley R. C. and Krueger R. F. (eds.), *Handbook of research methods in personality psychology*. The Guilford Press, London.
- Pelkonen, M., Marttunen, M., Pulkkinen, E., Laippala, P., Aro, H., 1997. Characteristics of out-patient adolescents with suicidal tendencies. *Acta Psychiatrica Scandinavica* 95, 100-107.
- Petronis, K.R., Samuels, J.F., Moscicki, E.K., Anthony, J.C., 1990. An epidemiologic investigation of potential risk factors for suicide attempts. *Social Psychiatry and Psychiatric Epidemiology* 25, 193-199.
- Placidi, G.P., Oquendo, M.A., Malone, K.M., Brodsky, B., Ellis, S.P., Mann, J.J., 2000. Anxiety in major depression: relationship to suicide attempts. *The American Journal of Psychiatry* 157, 1614-1618.
- Plener, P.L., Libal, G., Keller, F., Fegert, J.M., Muehlenkamp, J.J., 2009. An international comparison of adolescent non-suicidal self-injury (NSSI) and suicide attempts: Germany and the USA. *Psychological Medicine* 39, 1549-1558.
- Romans, S.E., Martin, J.L., Anderson, J.C., Herbison, G.P., Mullen, P.E., 1995. Sexual abuse in childhood and deliberate self-harm. *The American Journal of Psychiatry* 152, 1336-1342.
- Ross, S., Heath, N., 2002. A study of the frequency of self-mutilation in a community sample of adolescents. *Journal of Youth and Adolescence* 31, 67-77.
- Schmidtke, A., Bille-Brahe, U., De Leo, D., Kerkhof, A., 2004. Suicidal behaviour in Europe. Hogrefe & Huber, Gottingen.
- Schmidtke, A., Bille-Brahe, U., DeLeo, D., Kerkhof, A., Bjerke, T., Crepet, P., Haring, C., Hawton, K., Lonnqvist, J., Michel, K., Pommereau, X., Querejeta, I., Phillipe, I., Salander-Renberg, E., Temesvary, B., Wasserman, D., Fricke, S., Weinacker, B., Sampaio-Faria, J.G., 1996. Attempted suicide in Europe: rates, trends and sociodemographic characteristics of suicide attempters during the period 1989-1992. Results of the WHO/EURO Multicentre Study on Parasuicide. *Acta Psychiatrica Scandinavica* 93, 327-338.
- Schnyder, U., Valach, L., Bichsel, K., Michel, K., 1999. Attempted suicide. Do we understand the patients' reasons? *General Hospital Psychiatry* 21, 62-69.
- Shearer, S.L., Peters, C.P., Quaytman, M.S., Wadman, B.E., 1988. Intent and lethality of suicide attempts among female borderline inpatients. *American Journal of Psychiatry* 145, 1424-1427.
- Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, 59 Suppl 20, 22-33.

- Sinclair, J. M., Hawton, K., Gray, A., 2010. Six year follow-up of a clinical sample of self-harm patients. *Journal of affective disorder*, 121, 247-252.
- Skegg, K., 2005. Self-harm. *Lancet* 366, 1471-1483.
- Slee, N., Arensman, E., Garnefski, N., Spinhoven, P., 2007. Cognitive-behavioural therapy for deliberate self-harm. *Crisis* 28, 175-182.
- Slee, N., Garnefski, N., Spinhoven, P., Arensman, E., 2008a. The influence of cognitive emotion regulation strategies and depression severity on deliberate self-harm. *Suicide and Life Threatening Behaviour* 38, 274-286.
- Slee, N., Spinhoven, P., Garnefski, N., Arensman, E., 2008b. Emotion regulation as mediator of treatment outcome in therapy for deliberate self-harm. *Clinical Psychology and Psychotherapy* 15, 205-216.
- Suominen, K., Isometsa, E., Suokas, J., Haukka, J., Achte, K., Lonnqvist, J., 2004. Completed suicide after a suicide attempt: a 37-year follow-up study. *American Journal of Psychiatry* 161, 562-563.
- Suyemoto, K.L., 1998. The functions of self-mutilation. *Clinical Psychology Review* 18, 531-554.
- Taylor, R., Page, A., Morrell, S., Carter, G., Harrison, J., 2004. Socio-economic differentials in mental disorders and suicide attempts in Australia. *British Journal of Psychiatry* 185, 486-493.
- Ten Have, M., De Graaf, R., Van Dorsselaer, S., Verdurmen, J., Van 't, Land, H., Vollebergh, W., Beekman, A., 2009. Incidence and course of suicidal ideation and suicide attempts in the general population. *Canadian Journal of Psychiatry* 54, 824-833.
- Tuisku, V., Pelkonen, M., Karlsson, L., Kiviruusu, O., Holi, M., Ruuttu, T., Punamaki, R. L., Marttunen, M., 2006. Suicidal ideation, deliberate self-harm behaviour and suicide attempts among adolescent outpatients with depressive mood disorders and comorbid axis I disorders. *European Child and Adolescent Psychiatry* 15, 199-206.
- Tuisku, V., Pelkonen, M., Kiviruusu, O., Karlsson, L., Ruuttu, T., Marttunen, M., 2009. Factors associated with deliberate self-harm behaviour among depressed adolescent outpatients. *Journal of Adolescence* 32, 1125-1136.
- 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.
- Valentiner, D.P., Gutierrez, P.M., Blacker, D., 2002. Anxiety measures and their relationship to adolescent suicidal ideation and behaviour. *Journal of Anxiety Disorders* 16, 11-32.
- Van Vliet, I., De Beurs, E., 2007. [The MINI-International Neuropsychiatric Interview. A brief structured diagnostic psychiatric interview for DSM-IV en ICD-10 psychiatric disorders]. *Tijdschrift voor Psychiatrie* 49, 393-397.
- Weinberg, I., Ronningstam, E., Goldblatt, M.J., Schechter, M., Wheelis, J., Maltzberger, J.T., 2010. Strategies in treatment of suicidality: identification of common and treatment-specific interventions in empirically supported treatment manuals. *Journal of Clinical Psychiatry* 71, 699-706.
- Williams, J. M. G., Hassanyeh, F., 1983. Deliberate self-harm, clinical history and extreme scoring on the EPQ. *Personality and Individual Differences* 4, 347-350.

Chapter 5

Predicting outcome of depression using
the depressive symptom profile:
the Leiden Routine Outcome Monitoring Study

Martijn van Noorden
Esther van Fenema
Nic van der Wee
Frans Zitman
Erik Giltay

Depression and Anxiety (in press)

Abstract

Background: Being able to predict which patients are at risk for an unfavourable outcome is of high clinical importance. The aim of this study was to investigate the predictive value of items for individual depressive symptoms measured with the self-rated Beck Depression Inventory-Revised (BDI-II) self-report scale on outcome in a large naturalistic cohort of depressive outpatients.

Methods: We used a cohort of 1,489 adult patients aged 18-65 years with major depressive disorder or dysthymic disorder established with the MINI-Plus diagnostic interview. All patients had a Routine Outcome Monitoring (ROM) baseline measurement in 2004-2009, with a maximum of 2 years follow-up. We used multivariable Cox regression models to predict remission ($\text{MADRS} < 10$; where MADRS stands for Montgomery-Åsberg Depression Rating Scale) and response ($\geq 50\%$ improvement), and adjusted for clinical and demographic characteristics (i.e., marital status, level of education, working status, comorbid anxiety, avoidant and borderline personality traits, and suicidality) that were identified as predictors in earlier studies.

Results: Of the 21 BDI-II items, the items 'pessimism' and 'loss of energy' independently predicted for both remission and response. For pessimism, the HR for remission was 0.81 (95% confidence interval [CI]: 0.73-0.89, $p < 0.001$) and for loss of energy, the HR was 0.81 (0.72-0.92, $p = 0.001$).

Conclusions: These findings of robust prediction of poor outcome by baseline items of 'pessimism' and 'loss of energy' in a naturalistic treatment setting may help clinicians to identify depressive patients in need for additional or alternative therapeutic approaches.

Introduction

Despite extensive therapeutic options roughly one third of the patients with depressive disorders do not achieve remission after several adequate treatment trials (Rush et al., 2006), which may result in chronic depression (Rush et al., 1995). Being able to predict which patients are at risk for an unfavourable outcome is of high clinical importance, because these patients who do not achieve remission despite treatment may suffer considerably, and result in a poor cost-effectiveness of psychiatric care (Thase, 2003; Wittchen et al., 2011). In prospective studies, several factors have been associated with a poor outcome in patients with major depression. Sociodemographic factors, including low education (Barkow et al., 2003; McKenzie et al., 2010), unemployment, (Barkow et al., 2003; Sherbourne et al., 2004; Fournier et al., 2009; Frank et al., 2010) and being unmarried (Fournier et al., 2009; Frank et al., 2010; Meyers et al., 2002; Weinberger et al., 2008) were found to be independent predictors of a poor outcome. Also, clinical characteristics including greater severity of depressive symptoms (Barkow et al., 2003; Enns et al., 2005; Frank et al., 2010; Melartin et al., 2004; Meyers et al., 2002; Moos et al., 1999; Sargeant et al., 1990; Souery et al., 2007; Vuorilehto et al., 2009), loss of interest, reduced activity (Uher et al., 2011), suicidality (Barkow et al., 2003; Sherbourne et al., 2004; Souery et al., 2007), melancholic features (Souery et al., 2007), comorbid anxiety disorders (Enns & Cox, 2005; Melartin et al., 2004; Penninx et al., 2011; Souery et al., 2007; Weinberger et al., 2008) and personality traits or disorders such as avoidant, low extraversion and borderline (Enns & Cox, 2005; Skodol et al., 2011; Souery et al., 2007; Vuorilehto et al., 2009; Wiersma et al., 2011) were predictors of poor outcome. However, various definitions and operationalisations of poor outcome have been used in these studies, and duration of follow-up differed largely. Also, these studies have been conducted in different settings, e.g. community (Sargeant et al., 1990; Skodol et al., 2011), primary care (Barkow et al., 2003; Vuorilehto et al., 2009), psychiatric specialty care inpatient (McKenzie et al., 2010; Souery et al., 2007) and outpatient settings (Enns & Cox, 2005; Fournier et al., 2009; Frank et al., 2010; Meyers et al., 2002; Penninx et al., 2011; Souery et al., 2007; Uher et al., 2011; Weinberger et al., 2008). Moreover, most studies used data obtained from clinical trials, that limits generalisability to everyday clinical practice because of the usually much more extensive and stringent in- and exclusion criteria (van der Lem et al., 2011).

To our knowledge, no previous study investigated whether specific baseline symptoms, reflected in individual items on depressive rating scales at baseline predict outcome in a naturalistic clinical outpatient setting. If the presence of specific depressive symptoms at baseline would be of value in predicting outcome of depression, assessment

of these symptoms using a self-report measure could help clinicians to identify patients at risk of a poor outcome in a relatively easy way. Since the standardised assessment of symptom presence is increasingly becoming part of the usual clinical process, identification of symptoms that predict outcome could be helpful for clinician and patient.

A recent study from Uher et al. including 811 major depressive disorder (MDD) patients tested whether baseline depression symptom *dimensions* predicted outcome upon a 12-week-treatment with escitalopram or nortriptyline. The authors found and replicated that higher scores on the interest-activity dimension (i.e. low interest, reduced activity, indecisiveness and lack of enjoyment) predicted poorer treatment outcome (Uher et al., 2011).

The aim of the present study was to investigate whether individual baseline depressive symptoms measured with the widely used Beck Depression Inventory-Revised (BDI-II) self-report scale (Beck et al., 1988) would predict outcome in a large naturalistic cohort of depressive outpatients in a psychiatry specialty care setting. Outcome was defined as response and remission measured on the Montgomery-Åsberg Depression Rating Scale (MADRS) according to generally accepted criteria during up to 2 years of follow-up.

Methods

Participants

We used an initial cohort of 8,021 adult outpatients aged 18-65 years who were referred for treatment of a mood, anxiety, or somatoform (MAS) disorder to the Regional Mental Health Provider Rivierduinen (RD) and Leiden University Medical Center (LUMC). All patients had a Routine Outcome Monitoring (ROM) baseline assessment between 2004 and 2009, during intake before the start of treatment (de Beurs et al., 2011). 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 well-trained and supervised 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 (de Beurs et al., 2011). All questionnaires are completed on touch-screen computers, to prevent missing data within instruments. Patients with insufficient mastery of Dutch and patients unable to complete assessments are ineligible for ROM. On average, 80% of the referred patients were assessed with ROM in the study period. ROM data are primarily used for diagnosis and to inform clinicians and patients about treatment progress. The use of the anonymised data for research purposes has been approved by the Ethical Review Board of the LUMC.

Of these 8,021 patients, we selected all patients with a current MDD or dysthymic disorder according to the Mini International Neuropsychiatric Interview-Plus (MINI-Plus; $n=3,632$, 45.3%). The MINI-Plus is a standardised diagnostic interview that assesses current and lifetime Axis-I diagnoses according to the Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision (DSM-IV-TR; American Psychiatric Association, 2000). The inter-rater reliability ranges from 0.88 to 1.00, test-retest reliability from 0.76 to 0.93 and validity is adequate compared to the Composite International Diagnostic Interview (CIDI; Lecrubier, 1997; Sheehan et al., 1998; van Vliet et al., 2007). Patients with a lifetime bipolar disorder or primary psychotic illness were excluded, but a diagnosis of psychotic major depression according to the MINI-Plus was allowed. Of these 3,632 depressive patients, we selected all patients with a minimum of one follow-up ROM assessment ($n=1,988$, 54.7%). Patients who had only one ROM assessment were most likely not treated in the MAS outpatient clinics. Patients with a low baseline symptom severity (MADRS <10) were excluded ($n=87$, 2.4%) as well as patients with missing data on variables of interest ($n=412$, 11.3%). We thus included 1,489 patients in our analyses (41.0% of 3632 patients). There were no differences in age, gender, or level of education between included and excluded patients. Excluded patients had a lower MADRS baseline score than included patients (mean scores 21.9, standard deviation [SD] ± 9.7 vs. 24.1 ± 7.0 ; $t = -4.6$, $p < 0.001$).

Psychiatrists and clinical psychologists or psychotherapists in LUMC and RD provided outpatient treatment in accordance with the Dutch evidence based guidelines, consisting of pharmacotherapy, psychotherapy, or a combination of both. Treatment modalities were not taken into account in the analyses.

Measures

ROM assessments: We used the baseline ROM assessments and all successive ROM assessments with a maximum follow-up of 2 years. During the first ROM session the MINI-Plus was administered, as well as observer-rated and self-report scales, both generic and disorder-specific. In addition, demographic variables were collected (an overview of instruments is available at <http://www.lumc.nl/psychiatry/ROM-instruments>), as previously described in detail elsewhere (de Beurs et al., 2011; van Noorden et al., 2010).

Remission and response: Remission and response of depression were assessed using the MADRS. The MADRS is an observer-rated scale that assesses depression severity, and that is sensitive to change. The MADRS has an internal consistency (Cronbach's alpha) of 0.86, an inter-rater reliability coefficient of 0.65-0.97 (Montgomery, 1979). Remission was defined as the first follow-up ROM assessment at which the MADRS score of <10 had

been reached (Hawley et al., 2002; Zimmerman et al., 2004a; Zimmerman et al., 2004b) and response was defined as the first ROM assessment at which a reduction of $\geq 50\%$ of the baseline MADRS score had been achieved. The MADRS was administered at baseline and at every follow-up ROM assessment.

Individual depressive symptoms as possible predictors: Baseline depressive symptoms were measured with the BDI-II (Beck et al., 1996; Beck et al., 1988). The BDI-II is a 21-item self-report questionnaire that assesses the presence and severity of a broad range of depressive symptoms on a 4-point Likert-scale (0-3), with good psychometric properties (Beck et al., 1988). Symptoms measured with the BDI-II that most closely would fit in the interest/activity dimension found by Uher et al. (2008) are loss of pleasure, loss of interest, indecisiveness, loss of energy, concentration difficulties, and loss of interest in sex. In the present study, the BDI-II was chosen because of the broad range of symptom coverage in 21 individual items, and because of the practicability of a self-report measure. Furthermore, with the MADRS as relatively objective and 'gold standard' outcome measure, the risk of circularity as the result of functional relatedness of dependent and independent variables was likely to be lower than when the MADRS would have been used both as predictor and outcome (Senn, 1994).

Additional measures: 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 assessed, as well as housing situation, educational status and working situation. The Dimensional Assessment of Personality Pathology-Short Form (DAPP-SF) was administered at baseline to assess maladaptive personality traits (van Kampen et al., 2008). Additional scales in ROM that we used for this study were the observer-rated Brief Anxiety Scale (BAS; Tyrer et al., 1984), the patient-rated Brief Symptom Inventory (BSI; de Beurs et al., 2006; Derogatis et al., 1983), the observer-rated Clinical Global Impression (CGI; Guy, 1976) and the Global Assessment of Functioning (GAF) scale (American Psychiatric Association, 2000; Endicott et al., 1976).

Statistical analyses

Using descriptive statistics, baseline characteristics are described as number (percentage) or mean (SD, with interquartile range [IQR]), when appropriate. Univariable Hazard Ratios (HRs) of remission and response were computed according to baseline individual BDI-II items (predictor variables) using Cox regression on the time-to-remission and time-to-response endpoints. Possible collinearity of the individual BDI-II items was checked using Variance Inflation Factor (VIF) analysis. If the endpoint of remission or response was not reached within 2 years or the endpoint was not reached at the final follow-up

measurement, the case was censored at the last ROM assessment. All BDI-II items that predicted remission or response with HRs with p-values <0.10 in univariable analyses were subsequently selected for an initial forward stepwise multivariable Cox regression model, for both endpoints. We used criteria for selection and removal both of 0.10. In this multivariable model we adjusted for baseline MADRS score, age and gender. Furthermore, we repeated the multivariable model with adjustment for demographical factors (i.e., marital, educational and working status), comorbid anxiety (i.e., based on the MINI-Plus), avoidant and borderline personality traits (i.e., measured with the DAPP-SF), and current suicidal thoughts (i.e., based on the MINI-Plus) because these clinical characteristics had been associated with outcome in earlier studies (Barkow et al., 2003; Enns & Cox, 2005; Fournier et al., 2009; Frank et al., 2010; McKenzie et al., 2010; Melartin et al., 2004; Meyers et al., 2002; Moos & Cronkite, 1999; Penninx et al., 2011; Sargeant et al., 1990; Sherbourne et al., 2004; Skodol et al., 2011; Souery et al., 2007; Uher et al., 2011; Vuorilehto et al., 2009; Weinberger et al., 2008; Wiersma et al., 2011).

Based on the combined presence and or absence of the BDI-II symptoms 'pessimism' and 'loss of energy' four categories were constructed. In Kaplan-Meier analyses the cumulative incidence of remission is presented. Presence of a symptom was dichotomised as a score of 2 or 3 on the 4-point Likert scale.

We then performed sensitivity analyses in which we included only those patients with more severe complaints at baseline (MADRS ≥ 14), and additional analyses in which we excluded the patients with dysthymic disorder to include only patients with MDD. We used SPSS version 17.0 for all statistical analyses (SPSS Inc., Chicago, Ill).

Results

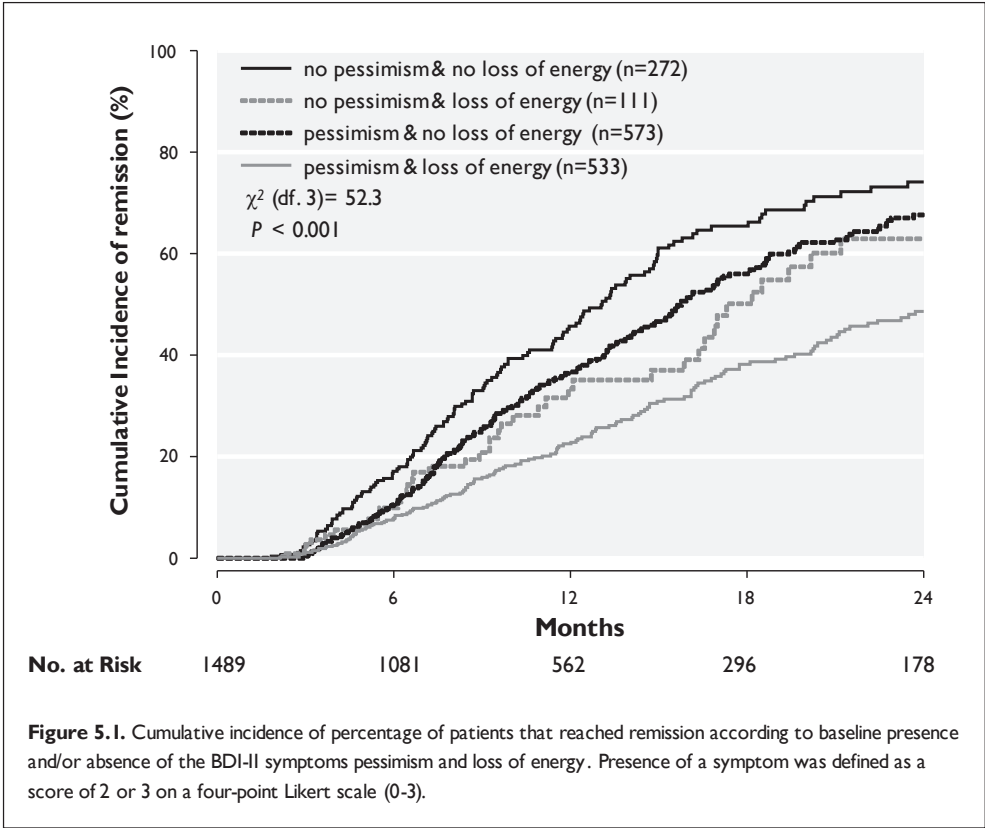
Baseline Characteristics

In the study sample of 1,489 depressed adults 62.3% were female. The mean age at ROM baseline assessment was 40.1 years (SD ± 12.1). Table 5.1 summarises the baseline characteristics of the sample. A Dutch ethnic background was found in 80.9%. The majority of the patients were living with a partner, 40.5% of the patients were currently working, 93.3% were diagnosed with MDD, and 6.7% with a dysthymic disorder. The mean MADRS score at baseline was 24.1 (SD ± 7.0), the mean BDI-II total score at baseline was 31.0 (SD ± 9.6), and the mean CGI score was 4.3 (SD ± 0.8).

Table 5.1. Baseline characteristics in 1,489 outpatients with MDD or Dysthymic Disorder

	n	%
Categorical baseline variables		
Female gender	928	62.3
Dutch ethnic background	1205	80.9
Marital status		
Married	799	53.7
Divorced/ widowed	261	17.5
Never married	429	28.8
Housing situation:		
Living alone	329	22.1
Living with partner	819	55.0
Living with family	341	22.9
Educational status:		
Lower education	636	42.7
Higher education	853	57.3
Employment situation:		
Working full-time	294	19.7
Working part time	309	20.8
Retired/unemployed	372	25.0
On sick leave	514	34.5
DSM-IV-TR diagnostic categories		
Major Depressive Disorder (MDD)	1389	93.3
Dysthymic Disorder (DD)	100	6.7
DSM-IV-TR comorbidity:		
Single Mood Disorder (MDD or DD)	705	47.3
Mood and Anxiety Disorder	577	38.8
Mood and Somatoform Disorder	122	8.2
Mood and Anxiety and Somatoform Disorder	85	5.7
Comorbidity of abuse or dependence :		
Alcohol abuse or dependence	96	6.4
Drug abuse or dependence	54	3.6
Continuous baseline variables		
	Mean (SD)	IQR
Age at interview (years)	40.1 (12.1)	30-50
MADRS score	24.1 (7.0)	19-28
BDI-II total score	31.0 (9.6)	24-38
BAS score	17.0 (6.1)	13-21
BSI total score	1.5 (0.7)	1.0-1.9
CGI score	4.3 (0.8)	4.0-5.0
GAF score	57.3 (6.8)	53-60

Abbreviations: SD: Standard Deviation; IQR: Interquartile Range; BSI: Brief Symptom Inventory; CGI: Clinical Global Impression; MADRS: Montgomery-Åsberg Depression Rating Scale; BDI-II: Beck Depression Inventory; BAS: Brief Anxiety Scale; GAF: Global Assessment of Functioning scale.



Predictors of remission

Within the follow-up time of 2 years, 562 patients (37.7%) achieved remission on the MADRS, defined as the first ROM assessment with a MADRS score <10 . In univariable Cox regression analyses 16 of 21 BDI items were associated with remission (Table 5.2). Among other symptoms, both ‘core symptoms of depression’, i.e. sadness and loss of interest, negatively predicted remission (HRs 0.76, 95% CI 0.68-0.86; $p<0.001$ and 0.81, 95% CI 0.74-0.89; $p<0.001$, respectively). In the first multivariable model, adjusted for age, gender and baseline MADRS score, only the BDI-II symptoms pessimism and loss of energy were independently associated with a lower chance of remission (HRs 0.81, 95% CI 0.73-0.89; $p<0.001$ and 0.81, 95% CI 0.72-0.92; $p=0.001$, respectively), and the core symptoms sadness and loss of interest were not. The additional multivariable model, adjusted for previously established clinical variables, yielded the same predictors (data not shown). Figure 5.1 shows the cumulative incidence of remission in Kaplan-Meier curves according to the combined items of pessimism and loss of energy. Presence of both pessimism and loss of energy predicted a 72.2% chance of

remission in 2 years, and absence of both symptoms predicted a 49.1% chance of remission in 2 years (HR 0.78, 95% CI 0.73-0.84; χ^2 52.3 (3); $p < 0.001$).

Table 5.2. Univariable and multivariable hazard ratios of remission on the MADRS according to baseline BDI -II scores in 1,489 outpatients with MDD or Dysthymic Disorder

BDI-II item	Remission MADRS Univariable model	p-value	Remission MADRS Multivariable model ^a	p-value
	HR (95% CI)		HR (95% CI)	
1 Sadness	0.76 (0.68-0.86)	<0.001		
2 Pessimism	0.73 (0.66-0.80)	<0.001	0.81 (0.73-0.89)	<0.001
3 Past failure	0.91 (0.84-0.99)	0.02		
4 Loss of pleasure	0.76 (0.68-0.85)	<0.001		
5 Feelings of guilt	0.90 (0.82-0.98)	0.02		
6 Punishment feelings	0.98 (0.91-1.05)	0.53		
7 Self dislike	0.95 (0.86-1.05)	0.29		
8 Self criticism	0.98 (0.90-1.06)	0.59		
9 Suicidal thoughts or wishes	0.81 (0.72-0.91)	<0.001		
10 Crying	0.98 (0.91-1.06)	0.64		
11 Agitation	0.97 (0.87-1.09)	0.61		
12 Loss of interest	0.81 (0.74-0.89)	<0.001		
13 Indecisiveness	0.84 (0.77-0.92)	<0.001		
14 Worthlessness	0.87 (0.80-0.95)	0.003		
15 Loss of energy	0.72 (0.64-0.80)	<0.001	0.81 (0.72-0.92)	0.001
16 Change of sleeping pattern	0.91 (0.83-1.00)	0.06		
17 Irritability	0.87 (0.79-0.95)	0.002		
18 Change of appetite	0.87 (0.80-0.95)	0.001		
19 Concentration difficulties	0.79 (0.71-0.88)	<0.001		
20 Tiredness or fatigue	0.80 (0.73-0.87)	<0.001		
21 Loss of interest in sex	0.86 (0.79-0.92)	<0.001		

Abbreviations: MADRS = Montgomery Åsberg Depression Rating Scale, BDI-II = Beck Depression Inventory, HR = Hazard Ratio, CI = Confidence Interval. Remission on the MADRS was defined as a score of < 10, response on the MADRS was defined as at least 50% reduction of the baseline score.

^a Multivariable models were adjusted for age, gender and baseline MADRS score.

Predictors of response

According to the criterion of at least 50% improvement on the MADRS score compared to baseline MADRS score, 650 patients (43.7%) were responder within 2 years. 10 of 21 BDI-II items were associated with treatment response in univariable analyses (Table 5.3). In the multivariable Cox regression model adjusted for age, gender and baseline MADRS score, again, only the BDI-II items pessimism and loss of energy independently predicted non-response with HRs of 0.86 (95% CI 0.78-0.94; $p = 0.001$) and 0.84 (95% CI 0.75-0.95; $p = 0.003$), respectively. Presence of both symptoms pessimism and loss of energy predicted a 61.1% chance of response in 2 years, and absence of both symptoms predicted a 49.4% chance of response in 2 years (HR 0.84, 95% CI 0.81-0.93; χ^2 20.2 (3); $p < 0.001$).

Sensitivity analyses

In the sensitivity analyses we first excluded patients with dysthymic disorder (n=100), leaving 1389 MDD patients. In the multivariable Cox regression model adjusted for age, gender and baseline MADRS score, results remained unchanged. Also, when we elevated the threshold for inclusion in analysis to a baseline MADRS score ≥ 14 the multivariable model in the 1410 remaining patients yielded the same results (data not shown).

Finally, the VIFs in a regression analysis were all < 2 , indicating that no multicollinearity existed between any of the BDI-II items.

Table 5.3. Univariable and multivariable hazard ratios of response on the MADRS according to baseline BDI -II scores in 1,489 outpatients with MDD or Dysthymic Disorder

BDI-II item	Response MADRS Univariable model		Response MADRS Multivariable model ^a	
	HR (95% CI)	p-value	HR (95% CI)	p-value
1 Sadness	0.88 (0.79-0.98)	0.02	0.86 (0.78-0.94)	0.001
2 Pessimism	0.83 (0.76-0.90)	< 0.001		
3 Past failure	0.97 (0.90-1.05)	0.50		
4 Loss of pleasure	0.86 (0.77-0.95)	0.004		
5 Feelings of guilt	0.96 (0.89-1.05)	0.39	0.84 (0.75-0.95)	0.003
6 Punishment feelings	1.02 (0.96-1.09)	0.25		
7 Self dislike	1.02 (0.93-1.12)	0.68		
8 Self criticism	1.04 (0.96-1.12)	0.39		
9 Suicidal thoughts or wishes	0.92 (0.82-1.02)	0.11		
10 Crying	0.99 (0.92-1.07)	0.85		
11 Agitation	1.05 (0.95-1.16)	0.38		
12 Loss of interest	0.90 (0.83-0.97)	0.007		
13 Indecisiveness	0.88 (0.82-0.95)	0.002		
14 Worthlessness	0.92 (0.85-1.00)	0.06		
15 Loss of energy	0.81 (0.73-0.90)	< 0.001		
16 Change of sleeping pattern	0.99 (0.91-1.08)	0.81		
17 Irritability	0.91 (0.84-0.99)	0.04		
18 Change of appetite	0.95 (0.88-1.03)	0.19		
19 Concentration difficulties	0.89 (0.80-0.98)	0.02		
20 Tiredness or fatigue	0.87 (0.79-0.94)	0.001		
21 Loss of interest in sex	0.91 (0.85-0.98)	0.009		

Abbreviations: MADRS: Montgomery-Åsberg Depression Rating Scale; BDI-II: Beck Depression Inventory-revised; HR: Hazard Ratio; CI: Confidence Interval. Remission on the MADRS was defined as a score of < 10 , response on the MADRS was defined as at least 50% reduction of the baseline score.

^a Multivariable models were adjusted for age, gender and baseline MADRS score.

Discussion

In this large naturalistic cohort of 1,489 depressive outpatients, of all 21 BDI-II items measured at baseline, pessimism and loss of energy independently predicted poor remission and response. These poor remission and response rates were found independently from characteristics that had been identified as outcome predictors in earlier studies (Barkow et al., 2003; Enns & Cox, 2005; Fournier et al., 2009; Frank et al., 2010; McKenzie et al., 2010; Melartin et al., 2004; Meyers et al. 2002; Moos & Cronkite, 1999; Penninx et al., 2011; Sargeant et al., 1990; Sherbourne et al., 2004; Skodol et al., 2011; Souery et al., 2007; Uher et al., 2011; Vuorilehto et al., 2009; Weinberger et al., 2008; Wiersma et al., 2011).

To our knowledge, this is the first large scale prospective study that investigated whether baseline individual depressive symptoms could predict outcome in a naturalistic sample of MDD patients. Our findings may facilitate early identification of patients at risk of poor outcome, in a relatively easy way by application of a self-report measure at baseline. This may have clinical consequences, since there is large consensus about remission of depression being the goal of treatment (Thase, 2003; Wade et al., 2009).

Our findings are mostly in line with the recent findings of Uher et al. (2011), who tested baseline symptom dimensions rather than items as predictors of treatment outcome. Based on a previous study of the same group (Uher et al., 2008), in which a factor analysis of the MADRS, BDI and the 17-item Hamilton Depression Rating Scale (HDRS) revealed three higher order factors (i.e. observed mood, cognitive and neurovegetative) that further split in six dimensions (i.e. mood and anxiety, pessimism and interest-activity, and sleep and appetite, respectively), the predictive value of those 9 symptom dimensions was tested. The interest-activity dimension was found to be the strongest predictor of treatment outcome. The individual symptoms 'pessimism' and 'lack of energy' are both symptoms within the 'cognitive' higher order factor as described by Uher et al. (2008), but our results suggest that in addition to the loss of energy/low interest/low activity factor, pessimism may be an important factor in predicting treatment outcome.

In earlier studies, pessimism was found to be a predictor of depression in the general population (Giltay et al., 2006), patients with breast cancer (Schou et al., 2004), and patients with urogenital cancer (Zenger et al., 2010), but these studies mainly studied low optimism or pessimism as dispositional traits, and patients were not diagnosed with MDD at baseline. Although dispositional optimism is not simply the reverse of pessimism, optimism and depression were investigated as predictors of physical and mental health functioning in a cohort of 659 veterans. Optimism was inversely correlated with depressive

symptoms, and depressive symptomatology was associated with reduced levels of psychosocial functioning. In this study, participants were not diagnosed with MDD either (Achat et al., 2000). Also, low dispositional optimism was associated with depression and a higher likelihood of starting psychotherapy in a large Finnish community study (Karlsson et al., 2011). In yet another study, pessimism was one of the predictors of suicide acts in 308 depressive patients that had been followed for 2 years (Oquendo et al., 2004). The strongest predictors of future suicidal acts were prior suicide attempt, severity of episode and cigarette smoking. In this study, however, treatment outcome was not the focus.

Loss of energy and fatigue, together with loss of interest or pleasure in usual activities were found to be predictors of chronicity in a study in 371 MDD patients that had been re-evaluated at 1, 4 and 10 years after initial screening and treatment (Moos and Cronkite, 1999). In this study by Moos et al. three symptom domains were constructed from a set of risk factors: severity of specific depressive symptoms, lack of self-confidence plus social isolation, and avoidance coping. However, all analyses were univariable and the authors stated that no follow-up diagnostic interviews had been conducted. Still, loss of energy was identified as a predictor of chronicity in this study as well. The authors hypothesise that loss of energy may be more persistent during treatment than depressed mood and therefore may have been a more robust predictor for outcome.

Our findings find support in the cognitive theory of depression (see for example Newman & Beck, 2009), in which dysfunctional thoughts and viewpoints are thought to be maintained by cognitive distortive processes and biases. Depressed patients may for example think of themselves as failures, and selectively absorb information that points in that direction, neglecting the evidence for the contrary. This mechanism is believed to result in deepening of pessimism and worsening of mood, resulting in a vicious circle of withdrawal from social interactions (Newman & Beck, 2009). One could hypothesise that pessimism leads to social withdrawal, and that increasing isolation results in loneliness (Rius-Ottenheim et al., 2011) and lowering of physical activities leading to and loss of energy. Another mechanism could be that pessimism/low optimism and loss of energy both negatively influence the motivation for treatment resulting in lower adherence to treatment protocols (Leedham et al., 1995). Since the BDI-II measures state rather than trait symptoms, interpretations with respect to dispositional optimism/pessimism are not possible with the present results.

Our study has several strengths. First, this study was conducted in a large naturalistic sample of depressive patients increasing the generalisability of findings. Second, the extensive standard battery of our ROM assessments included a structured clinical interview, the BDI-II and the observer-rated scale (MADRS). Third, outcome

predictors identified in earlier studies were taken into account in our multivariable analyses. Fourth, assessments were done by well-trained research nurses who were not involved in treatment.

Possible limitations of our study include the absence of information regarding treatment in ROM. We assume that patients were mostly 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 (van der Lem et al., 2011). Second, no information on psychiatric history, family history and somatic comorbidity of the patients was available for our analyses. These parameters are assessed by clinicians and are not part of ROM. Third, the large proportion of patients lost to follow up at each ROM assessment is a limitation of our study. However, we believe that this large attrition rate is inherent to the design of the study that solely relies on ROM data collected as part of naturalistic treatment. A considerable loss to follow-up is common in cohort studies and trials with a 'real world' approach. In step II of the STAR*D multicenter trial for example, a study that aimed to be represent an everyday treatment setting, the loss to follow-up was 30% (Rush et al., 2006). 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 conclusion, we found that the presence of baseline symptoms pessimism and lack of energy measured with the BDI-II strongly predicted naturalistic outcome in MDD outpatients. Systematic assessment of these depressive symptoms using validated rating scales is relatively easy to implement and often already form an invaluable part of clinical practice. Besides the goal to more objectively inform clinicians and patients about treatment progress, it may help the clinician in risk assessment of MDD patients already during the initial assessment of the intake phase.

References

- Achat, H., Kawachi, I., Spiro, A., DeMolles, D.A., Sparrow, D., 2000. Optimism and depression as predictors of physical and mental health functioning: the Normative Aging Study. *Annals of Behavioral Medicine* 22, 127-130.
- American Psychiatric Association, 2000. *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC.
- Barkow, K., Maier, W., Ustun, T.B., Gansicke, M., Wittchen, H.U., Heun, R., 2003. Risk factors for depression at 12-month follow-up in adult primary health care patients with major depression: an international prospective study. *Journal of Affective Disorders* 76, 157-169.
- Beck, A.T., Steer, R.A., Brown, O.K., 1996. *Beck Depression Inventory Manual*. Psychological Corporation, San Antonio, TX.
- Beck, A.T., Steer, R.A., Carbin, M.G., 1988. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review* 8, 77-100.
- de Beurs, E., den Hollander-Gijsman, M.E., van Rood, Y.R., van der Wee, N.J., Giltay, E.J., van Noorden, M.S., van der Lem, R., van Fenema, E.M., 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 and Psychotherapy* 18, 1-12.
- de Beurs, E., Zitman, F.G., 2006. The Brief Symptom Inventory (BSI): Reliability and validity of a practical alternative to SCL-90. *Maandblad Geestelijke Volksgezondheid* 61, 120-141.
- Derogatis, L.R., Melisaratos, N., 1983. The Brief Symptom Inventory: an introductory report. *Psychological Medicine* 13, 595-605.
- Endicott, J., Spitzer, R.L., Fleiss, J.L., Cohen, J., 1976. The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry* 33, 766-771.
- Enns, M.W., Cox, B.J., 2005. Psychosocial and clinical predictors of symptom persistence vs remission in major depressive disorder. *Canadian Journal of Psychiatry* 50, 769-777.
- Fournier, J.C., DeRubeis, R.J., Shelton, R.C., Hollon, S.D., Amsterdam, J.D., Gallop, R., 2009. Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression. *Journal of Consulting and Clinical Psychology* 77, 775-787.
- Frank, E., Cassano, G.B., Rucci, P., Thompson, W.K., Kraemer, H.C., Fagiolini, A., Maggi, L., Kupfer, D.J., Shear, M.K., Houck, P.R., Calugi, S., Grochocinski, V.J., Scocco, P., Buttenfield, J., Forgione, R.N., 2011. Predictors and moderators of time to remission of major depression with interpersonal psychotherapy and SSRI pharmacotherapy. *Psychological Medicine* 41, 151-162.
- Giltay, E.J., Zitman, F.G., Kromhout, D., 2006. Dispositional optimism and the risk of depressive symptoms during 15 years of follow-up: the Zutphen Elderly Study. *Journal of Affective Disorders* 91, 45-52.
- Guy, W., 1976. *ECDEU Assessment Manual for Psychopharmacology*. National Institute of Mental Health, Rockville, MD.
- Hawley, C.J., Gale, T.M., Sivakumaran, T., 2002. Defining remission by cut off score on the MADRS: selecting the optimal value. *Journal of Affective Disorders* 72, 177-184.

- 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.
- Karlsson, H., Kronstrom, K., Nabi, H., Oksanen, T., Salo, P., Virtanen, M., Suominen, S., Kivimaki, M., Vahtera, J., 2011. Low level of optimism predicts initiation of psychotherapy for depression: results from the Finnish Public Sector Study. *Psychotherapy and Psychosomatics* 80, 238-244.
- Lecrubier, Y., 1997. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *European psychiatry* 12, 224-231.
- Leedham, B., Meyerowitz, B.E., Muirhead, J., Frist, W.H., 1995. Positive expectations predict health after heart transplantation. *Health Psychology* 14, 74-79.
- van der Lem, R., van der Wee, N.J., van Veen, T., Zitman, F.G., 2011. The generalizability of antidepressant efficacy trials to routine psychiatric out-patient practice. *Psychological Medicine* 41, 1353-1363.
- McKenzie, M., Clarke, D.M., McKenzie, D.P., Smith, G.C., 2010. Which factors predict the persistence of DSM-IV depression, anxiety, and somatoform disorders in the medically ill three months post hospital discharge? *Journal of Psychosomatic Research* 68, 21-28.
- Melartin, T.K., Rytala, H.J., Leskela, U.S., Lestela-Mielonen, P.S., Sokero, T.P., Isometsa, E.T., 2004. Severity and comorbidity predict episode duration and recurrence of DSM-IV major depressive disorder. *Journal of Clinical Psychiatry* 65, 810-819.
- Meyers, B.S., Sirey, J.A., Bruce, M., Hamilton, M., Raue, P., Friedman, S.J., Rickey, C., Kakuma, T., Carroll, M.K., Kiosses, D., Alexopoulos, G., 2002. Predictors of early recovery from major depression among persons admitted to community-based clinics: an observational study. *Archives of General Psychiatry* 59, 729-735.
- Montgomery, S.A., 1979. A new depression scale designed to be sensitive to change. *The British journal of psychiatry* 134, 382-389.
- Moos, R.H., Cronkite, R.C., 1999. Symptom-based predictors of a 10-year chronic course of treated depression. *Journal of Nervous and Mental Disorders* 187, 360-368.
- Newman, C.F., Beck, A.T., 2009. Cognitive therapy. In: Sadock, B.J., Sadock, V.A. (eds.), *Kaplan and Sadock's Comprehensive Textbook of Psychiatry*. Lippincott Williams & Wilkins, Baltimore.
- van Noorden, M.S., Giltay, E.J., den Hollander-Gijsman, M.E., van der Wee, N.J., 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.
- Oquendo, M.A., Galfalvy, H., Russo, S., Ellis, S.P., Grunebaum, M.F., Burke, A., Mann, J.J., 2004. Prospective study of clinical predictors of suicidal acts after a major depressive episode in patients with major depressive disorder or bipolar disorder. *The American Journal of Psychiatry* 161, 1433-1441.
- Penninx, B.W., Nolen, W.A., Lamers, F., Zitman, F.G., Smit, J.H., Spinhoven, P., Cuijpers, P., de Jong, P.J., van Marwijk, H.W., van der Meer, K., Verhaak, P., Laurant, M.G., de Graaf, R., Hoogendijk, W.J., van der Wee, N.J., Ormel, J., van Dyck, R., Beekman, A.T., 2011. Two-year course of depressive and anxiety disorders: Results from the Netherlands Study of Depression and Anxiety (NESDA). *Journal of Affective Disorders* (in press).

- Rius-Ottenheim, N., Kromhout, D., van der Mast, R.C., Zitman, F.G., Geleijnse, J.M., Giltay, E.J., 2011. Dispositional optimism and loneliness in older men. *International Journal of Geriatric Psychiatry* (in press).
- Rush, A.J., Laux, G., Giles, D.E., Jarrett, R.B., Weissenburger, J., Feldman-Koffler, F., Stone, L., 1995. Clinical characteristics of outpatients with chronic major depression. *Journal of Affective Disorders* 34, 25-32.
- Rush, A.J., Trivedi, M.H., Wisniewski, S.R., Nierenberg, A.A., Stewart, J.W., Warden, D., Niederehe, G., Thase, M.E., Lavori, P.W., Lebowitz, B.D., McGrath, P.J., Rosenbaum, J.F., Sackeim, H.A., Kupfer, D.J., Luther, J., Fava, M., 2006. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *The American Journal of Psychiatry* 163, 1905-1917.
- Sargeant, J.K., Bruce, M.L., Florio, L.P., Weissman, M.M., 1990. Factors associated with 1-year outcome of major depression in the community. *Archives of General Psychiatry* 47, 519-526.
- Schou, I., Ekeberg, O., Ruland, C.M., Sandvik, L., Karesen, R., 2004. Pessimism as a predictor of emotional morbidity one year following breast cancer surgery. *Psychooncology* 13, 309-320.
- Senn, S., 1994. Importance of trends in the interpretation of an over-all odds ratio in the meta-analysis of clinical trials. *Statistics in Medicine* 13, 293-296.
- Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry* 59 Suppl 20, 22-33.
- Sherbourne, C., Schoenbaum, M., Wells, K.B., Croghan, T.W., 2004. Characteristics, treatment patterns, and outcomes of persistent depression despite treatment in primary care. *General Hospital Psychiatry* 26, 106-114.
- Skodol, A.E., Grilo, C.M., Keyes, K.M., Geier, T., Grant, B.F., Hasin, D.S., 2011. Relationship of personality disorders to the course of major depressive disorder in a nationally representative sample. *The American Journal of Psychiatry* 168, 257-264.
- Souery, D., Oswald, P., Massat, I., Bailer, U., Bollen, J., Demyttenaere, K., Kasper, S., Lecrubier, Y., Montgomery, S., Serretti, A., Zohar, J., Mendlewicz, J., 2007. Clinical factors associated with treatment resistance in major depressive disorder: results from a European multicenter study. *Journal of Clinical Psychiatry* 68, 1062-1070.
- Thase, M.E., 2003. Achieving remission and managing relapse in depression. *Journal of Clinical Psychiatry* 64 Suppl 18, 3-7.
- 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.
- Uher, R., Farmer, A., Maier, W., Rietschel, M., Hauser, J., Marusic, A., Mors, O., Elkin, A., Williamson, R.J., Schmael, C., Henigsberg, N., Perez, J., Mendlewicz, J., Janzing, J.G., Zobel, A., Skibinska, M., Kozel, D., Stamp, A.S., Bajs, M., Placentino, A., Barreto, M., McGuffin, P., Aitchison, K.J., 2008. Measuring depression: comparison and integration of three scales in the GENDEP study. *Psychological Medicine* 38, 289-300.
- Uher, R., Perlis, R.H., Henigsberg, N., Zobel, A., Rietschel, M., Mors, O., Hauser, J., Dernovsek, M.Z., Souery, D., Bajs, M., Maier, W., Aitchison, K.J., Farmer, A., McGuffin, P., 2011. Depression symptom dimensions as predictors of antidepressant treatment outcome: replicable evidence for interest-activity symptoms. *Psychological Medicine* (in press).

- van Vliet, I.M., de Beurs, E., 2007. [The MINI-International Neuropsychiatric Interview. A brief structured diagnostic psychiatric interview for DSM-IV en ICD-10 psychiatric disorders]. *Tijdschrift voor Psychiatrie* 49, 393-397.
- Vuorilehto, M.S., Melartin, T.K., Isometsä, E.T., 2009. Course and outcome of depressive disorders in primary care: a prospective 18-month study. *Psychological Medicine* 39, 1697-1707.
- Wade, A.G., Schlaepfer, T.E., Andersen, H.F., Kilts, C.D., 2009. Clinical milestones predict symptom remission over 6-month and choice of treatment of patients with major depressive disorder (MDD). *Journal of Psychiatry Research* 43, 586-575.
- Weinberger, M.I., Sirey, J.A., Bruce, M.L., Heo, M., Papademetriou, E., Meyers, B.S., 2008. Predictors of major depression six months after admission for outpatient treatment. *Psychiatric Services* 59, 1211-1215.
- Wiersma, J.E., Van Oppen, P., Van Schaik, D.J., van der Does, A.J., Beekman, A.T., Penninx, B.W., 2011. Psychological characteristics of chronic depression: a longitudinal cohort study. *Journal of Clinical Psychiatry* 72, 288-294.
- Wittchen, H.U., Jacobi, F., Rehm, J., Gustavsson, A., Svensson, M., Jönsson, B., Olesen, J., Allgulander, C., Alonso, J., Faravelli, C., Fratiglioni, L., Jennum, P., Lieb, R., Maercker, A., van Os, J., Preisig, M., Salvador-Carulla, L., Simon, R., Steinhausen, H.C., 2011. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European Neuropsychopharmacology* 21, 655-679.
- Zenger, M., Brix, C., Borowski, J., Stolzenburg, J.U., Hinz, A., 2010. The impact of optimism on anxiety, depression and quality of life in urogenital cancer patients. *Psychooncology* 19, 879-886.
- Zimmerman, M., Chelminski, I., Young, D., 2004a. On the threshold of disorder: a study of the impact of the DSM-IV clinical significance criterion on diagnosing depressive and anxiety disorders in clinical practice. *Journal of Clinical Psychiatry* 65, 1400-1405.
- Zimmerman, M., Posternak, M.A., Chelminski, I., 2004b. Derivation of a definition of remission on the Montgomery-Åsberg depression rating scale corresponding to the definition of remission on the Hamilton rating scale for depression. *Journal of Psychiatric Research* 38, 577-582.

Chapter 6

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

Journal of Affective Disorders (in press)

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: We conducted a cohort study in 892 adult MAS patients in a naturalistic outpatient psychiatric specialty care setting and validated our results in a replication cohort of 1,392 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 Impression-Severity Scale (CGI-S), during follow-up during up to 2 years.

Results: In multivariable Cox regression models, independent and replicated predictors for poor outcome were higher age (overall $p < 0.001$ for combined cohorts in multivariable Cox regression model), having comorbid MAS disorders or a somatoform disorder ($p < 0.001$), dysfunctional personality traits (i.e., tendency to self-harm [$p < 0.001$], intimacy problems [$p < 0.001$], affective lability [$p < 0.001$]), and a low reported general health status ($p < 0.001$).

Conclusions: MAS patients meeting the profile of being elderly, suffering from comorbid MAS disorders or a somatoform disorder, with cluster B personality traits, and a poor reported general health may need special preventive measures to minimise 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 (Wittchen et al., 2005). 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 (de Waal et al., 2004; Gaynes, 2009; Kessler et al., 2005; Maier et al., 1999). 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 (Fournier et al., 2009; Frank et al., 2011; Meyers et al., 2002; Weinberger et al., 2008), being unemployed (Fournier et al., 2009; Frank et al., 2011), a lower level of education (McKenzie et al., 2010), a greater severity of depressive symptoms (Frank et al., 2011; Melartin et al., 2004; Meyers et al., 2002; Souery et al., 2007), concomitant symptoms of pain, comorbid anxiety disorders (Melartin et al., 2004; Penninx et al., 2011; Souery et al., 2007; Weinberger et al., 2008) and borderline personality disorder (Skodol et al., 2011; Souery et al., 2007). Poorer remission rates from anxiety disorders were predicted by a higher severity of the anxiety symptoms, concomitant symptoms of pain, comorbid depression (Penninx et al., 2011) and prevalent personality disorders (Massion et al., 2002; Yonkers et al., 2000). Predictors of low remission rates for somatoform disorders were a lower level of education (McKenzie et al., 2010), concomitant symptoms of depression (Leiknes et al., 2007), a greater severity of the somatoform symptoms, and suffering from a comorbid personality disorder (Phillips et al., 2005).

However, the available studies did not include patients with different MAS diagnoses concomitantly and showed important methodological differences. Sample size ranging from 165 (Meyers et al., 2002) to 1,996 (Skodol et al., 2011), a duration of follow-up ranging from 3 months (Frank et al., 2011) to 5 years (Massion et al., 2002), the use of samples from highly selected populations (Fournier et al., 2009; Frank et al., 2011), 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 initialised, repeatedly every 3-4 months during treatment (de Beurs et al., 2011; van Noorden et al., 2010). 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 (de Beurs et al., 2011). 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 standardised 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 (MINI-Plus; Sheehan et al., 1998). The MINI-Plus has good psychometric properties, with inter-rater reliability ranging from 0.88-1.00, test-retest reliability ranging from 0.76-0.93 and adequate validity compared to the Composite International Diagnostic Interview (Lecrubier, 1997).

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 Improvement Scale (CGI). The BSI is a 53-item version of

the Symptom Checklist (SCL-90; Derogatis et al., 1983) 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 0.71-0.85 (de Beurs et al., 2011; Derogatis & Melisaratos, 1983) and test-retest reliability coefficients ranging from 0.68-0.91. The convergent validity has proven to be very good (Derogatis & Melisaratos, 1983). The CGI is a simple standardised observer-rated assessment tool for making global assessments of the severity of illness (Guy, 1976). 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'; CGI-S), was used in the present analyses. The CGI is widely used in medical care and clinical research as an outcome measure because of its face validity and practicability (Kadouri et al., 2007).

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 (Roy-Byrne et al., 2010) 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 2,876 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 2,966 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 (Roy-Byrne et al., 2010). Then we excluded all patients who did not have follow-up ROM assessments after

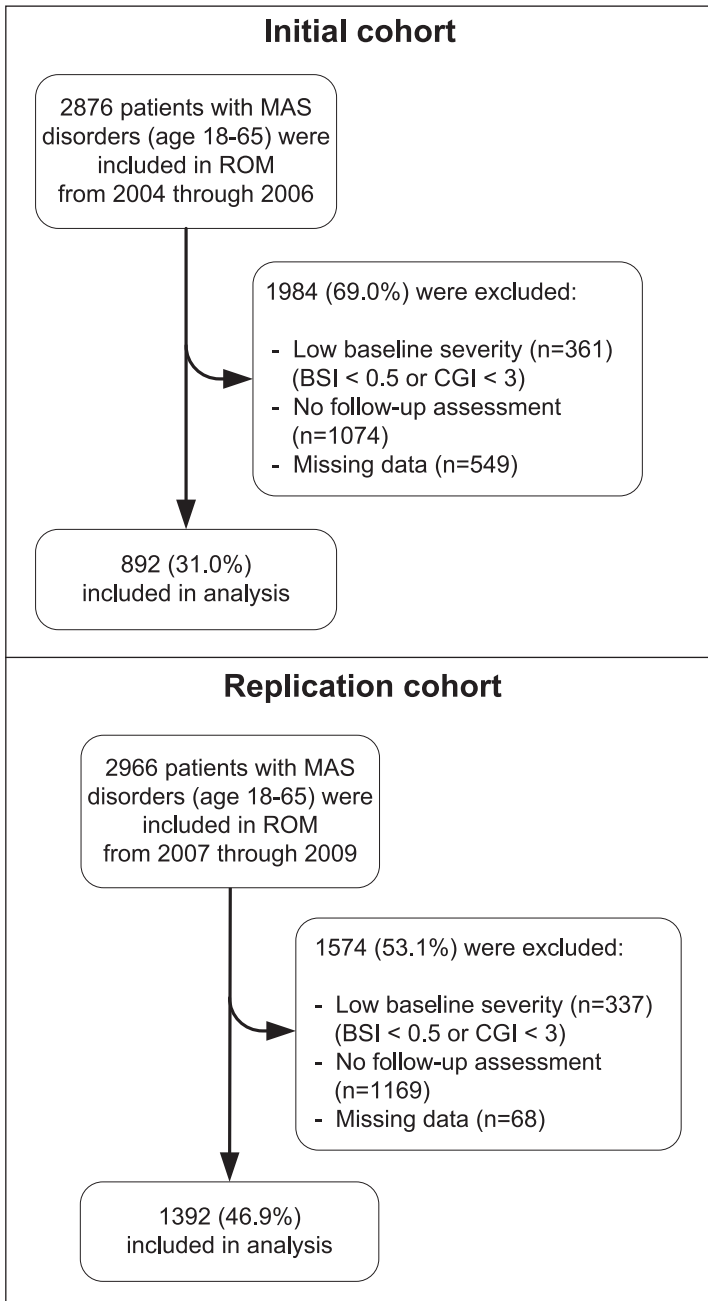


Figure 6.1. Flowchart of the patients in the initial and replication cohorts

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 2,876 patients (31.0%) were included, and in the replication cohort 1,392 of 2,966 patients (46.9%; Figure 6.1). 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%, $\chi^2(2) = 252.43$, $p < 0.001$). 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 (Van Fenema et al., 2010). 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 categorised in 'married' (which also included living together in a relationship), 'divorced or widowed', and 'never married'. Housing situation was categorised 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 categorised 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 categorised as: mood disorders (major depressive disorder or dysthymic disorder), anxiety disorders (panic disorder, social phobia, generalised anxiety disorder or posttraumatic stress 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 one 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-Åsberg Depression Rating Scale (MADRS), the Brief Anxiety Scale (BAS) and a scale that assesses psychomotor inhibition (INH; Goekoop et al., 1992). The MADRS has an internal consistency (Cronbach's alpha) of 0.86, and an inter-rater reliability coefficient of 0.65-0.97 (Montgomery et al., 1979).

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 (Vatnaland et al., 2007).

All participants also completed the Mood and Anxiety Symptom Questionnaire (MASQ), a questionnaire that measures the dimensions of the tripartite model of anxiety and depression (Watson et al., 1995). 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 0.91-0.96 across subscales (Watson et al., 1995).

The Dimensional Assessment of Personality Pathology-Short Form (DAPP-SF) was administered to assess maladaptive personality traits (van Kampen et al., 2008). 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 0.78-0.89 across subscales (van Kampen et al., 2008).

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 (Ware & Sherbourne, 1992). The SF-36 has good psychometric properties with Cronbach's alphas ranging from 0.66-0.94 across subscales (Ware & Sherbourne, 1992).

Statistical analyses

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 6.1 shows the predictor variables and DSM-IV-TR diagnoses at baseline of 892 and 1,392 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.

Univariable predictors of poor outcome

Table 6.2 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 6.3 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.

Table 6.1. Baseline characteristics in 892 and 1,392 outpatients with MAS disorders

	Initial cohort (n=892)		Replication cohort (n=1392)	
	n	%	n	%
Categorical predictor variables				
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 ($\pm$ SD)	IQR	Mean ($\pm$ 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-S score	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

Abbreviations: MAS: Mood, Anxiety and Somatoform; SD: Standard Deviation; IQR: Interquartile Range; BSI: Brief Symptom Inventory; CGI-S: Clinical Global Impression-Severity; MADRS: Montgomery-Åsberg Depression Rating Scale; BAS: Brief Anxiety Scale; GAF: Global Assessment of Functioning.

Table 6.2. Univariable hazard ratios of poor outcome according to baseline categorical predictor variables in 892 outpatients from the initial cohort

Categorical predictor variables	BSI poor outcome HR (95% CI)	p-value	CGI-S poor outcome HR (95% CI)	p-value	BSI or CGI-S poor outcome 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
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

Abbreviations: BSI: Brief Symptom Inventory; CGI-S: Clinical Global Impression; CI: Confidence Interval; HR: Hazard Ratio; DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision; Ref: 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-S was defined as a score of ≥ 3 on the 7-point Likert scale after up to 24 months fu.

Table 6.3. Univariable hazard ratios of poor outcome according to baseline continuous predictor variables in 892 outpatients from the initial cohort

Continuous predictor variables	BSI poor outcome HR (95% CI)	p-value	CG-S poor outcome HR (95% CI)	p-value	BSI or CG-S poor outcome 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
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
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

Abbreviations: BSI: Brief Symptom Inventory; CGI-S: Clinical Global Impression-Severity; CI: Confidence Interval; HR: Hazard Ratio; MADRS: Montgomery-Åsberg Depression Rating Scale; BAS: Brief Anxiety Scale; GAF: Global Assessment of Functioning; MAS: Mood, Anxiety and Somatoform; MASQ: Mood and Anxiety Symptoms Questionnaire; DAPP-SF: Dimensional Assessment of Personality Pathology-Short Form; SF-36: 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.

Independent predictors of poor outcome

Table 6.4 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.

Table 6.4. Multivariable Hazard Ratios of poor outcome in 892 and 1,392 outpatients with MAS disorders

Variables	Initial cohort (n=892) HR (95% CI)	p-value	Replication cohort (n=1,392) 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 Liability	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

Abbreviations: MAS: Mood, Anxiety and Somatoform; HR: Hazard Ratio; CI: Confidence Interval; DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision; Ref: Reference group; DAPP-SF: Dimensional Assessment of Personality Pathology short -form; SF-36: Short Form 36. Poor outcome was defined as <50% reduction of the baseline BSI score or a 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 initial 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.

Next, a 1-SD increase in the DAPP-SF Affective Liability subscale was associated with an 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. unfavourable scores on the General Health subscale that predicted 22% and 26% higher chances of poor outcome, respectively.

Figure 6.2 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 1,392 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 (Fournier et al., 2009) and outpatients with depression and/or anxiety (Penninx et al., 2011). Physical illness, disability, social isolation and loneliness may become more common with advancing age, which may adversely affect treatment outcome of comorbid psychiatric disorders (Luanaigh et al., 2008). Second, comorbidity has repeatedly been related to poor outcomes (de Waal et al., 2004; Frank et al., 2011; Leiknes et al., 2007; Maier & Falkai, 1999; Melartin et al., 2004; Penninx et al., 2011; Souery et al., 2007), which may be due to the higher burden of disease and severity (de Waal et al., 2004; Maier & Falkai, 1999). Third, somatic symptoms and somatic comorbidity have been linked to poor outcomes (McKenzie et al., 2010; Speckens et al., 1996), 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 (Massion et al., 2002; Meyers et al., 2002; Phillips et al.,

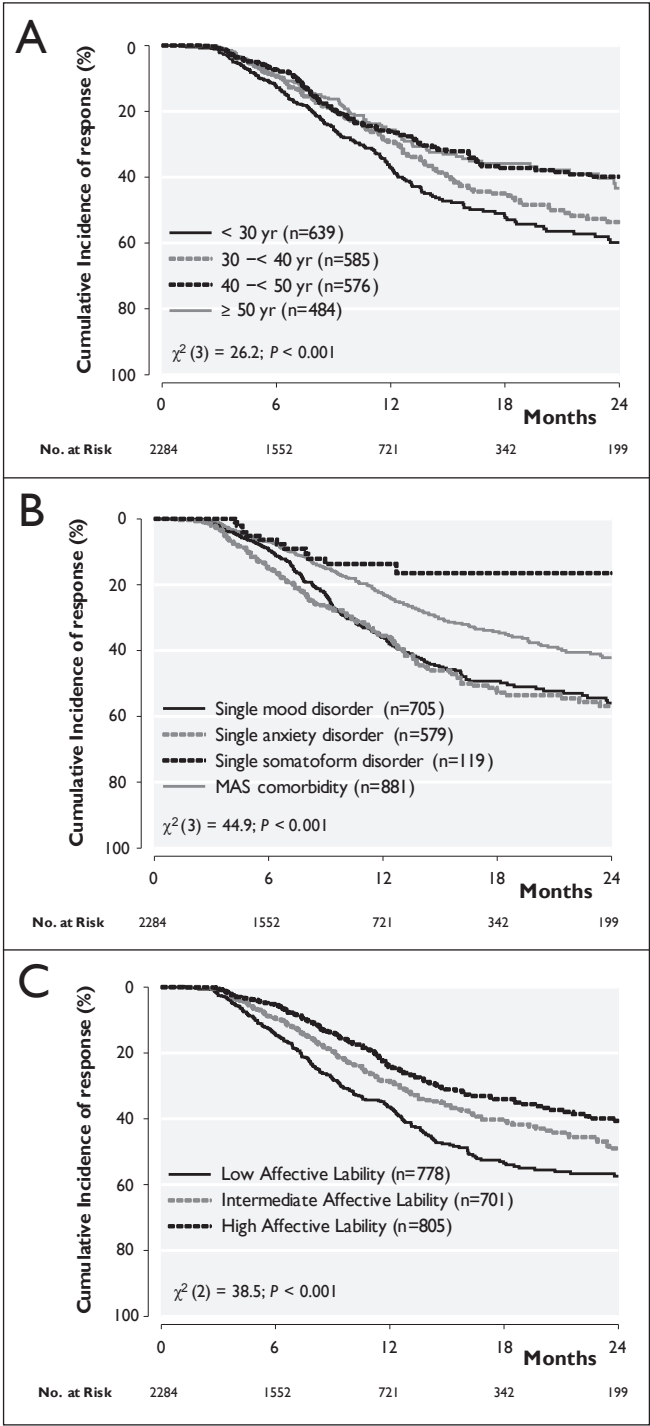
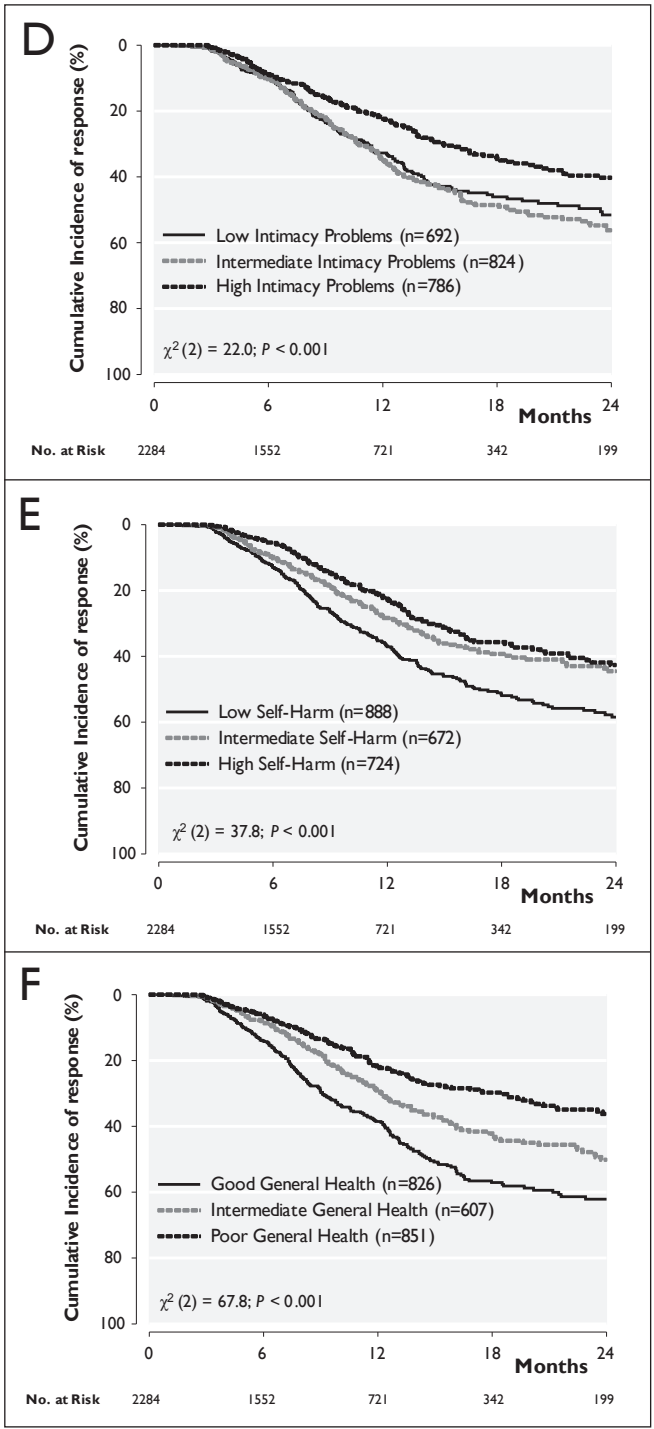


Figure 6.2. Kaplan-Meier curves for poor outcome according to baseline characteristics; Kaplan-Meier curves are shown for the incidence of naturalistic treatment response based on a >50% improvement on the BSI and a score of 1 or 2 on the CGI-S among 2,284 patients with a MAS disorder from the initial and replication cohorts combined. Patients were compared for age per decade (Box A), DSM-IV-TR diagnostic category (Box B), Affective Liability (Box C), Intimacy Problems (Box D), Self-Harm (Box E), and General Health (Box F). To facilitate graphical presentation of data, categorisation of Affective Liability, Intimacy Problems, Self-Harm and General Health into tertiles was applied.



2005; Skodol et al., 2011; Souery et al., 2007; Yonkers et al., 2000). Personality disorders at baseline in a sample of 303 patients with a depressive disorder were associated with long-term adverse outcomes (Grilo et al., 2010). In a study among 514 patients with generalised 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 (Massion et al., 2002). Finally, somatic conditions and poor general health have often been associated with depressive and anxiety disorders (Katon et al., 2010; McKenzie et al., 2010; Mezuk et al., 2008; Weinberger et al., 2008). 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 behaviour, predicted poor outcome and should be a focus of attention during treatment (van Kampen et al., 2008). MAS patients with more pronounced personality traits may benefit from specific therapeutic approaches, such as targeting cluster B personality characteristics, or cognitive behavioural therapy or dialectical behaviour therapy that combines cognitive-behavioural techniques with concepts of distress tolerance, acceptance, and mindful awareness (Robins et al., 2004). 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 behaviour, 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 (Katon et al., 2010). 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-S are clinically useful instruments to evaluate treatment outcomes across the diverse MAS disorders and to assess the patient population in its entirety (Klinkman, 2009; Trivedi, 2001). Finally, all assessments were made by trained research nurses who were not involved in treatment.

Our study also has several potential limitations. First, 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 (van der Lem et al., 2011; Van Fenema et al., 2010). Second, no information on psychiatric history, family history and somatic comorbidity of the patients was available for our cohorts. Third, there was substantial attrition, as many patients completed only one ROM assessment. A considerable loss to follow-up is common in 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 (Rush et al., 2006). Third, 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. Fourth, the BSI may not be specific enough to fully capture clinical changes in all MAS disorders, especially in somatoform disorders (Arnold et al., 2010). 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.

References

- Arnold, L.M., Leon, T., Whalen, E., Barrett, J., 2010. Relationships among pain and depressive and anxiety symptoms in clinical trials of pregabalin in fibromyalgia. *Psychosomatics* 51, 489-497.
- de Beurs, E., den Hollander-Gijsman, M.E., van Rood, Y.R., van der Wee, N.J., Giltay, E.J., van Noorden, M.S., van der Lem, R., van Fenema, E.M., 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 and Psychotherapy* 18, 1-12.
- Derogatis, L.R., Melisaratos, N., 1983. The Brief Symptom Inventory: an introductory report. *Psychological Medicine* 13, 595-605.
- van Fenema, E.M., van der Wee, N.J., Giltay, E.J., den Hollander-Gijsman, M.E., Zitman, F.G., 2010. Vitality predicts level of guideline-concordant care in routine treatment of mood, anxiety and somatoform disorders. *Journal of Evaluation in Clinical Practice* (in press).
- Fournier, J.C., DeRubeis, R.J., Shelton, R.C., Hollon, S.D., Amsterdam, J.D., Gallop, R., 2009. Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression. *Journal of Consulting and Clinical Psychology* 77, 775-787.
- Frank, E., Cassano, G.B., Rucci, P., Thompson, W.K., Kraemer, H.C., Fagiolini, A., Maggi, L., Kupfer, D.J., Shear, M.K., Houck, P.R., Calugi, S., Grochocinski, V.J., Scocco, P., Buttenfield, J., Forgiione, R.N., 2011. Predictors and moderators of time to remission of major depression with interpersonal psychotherapy and SSRI pharmacotherapy. *Psychological Medicine* 41, 151-162.
- Gaynes, B.N., 2009. Identifying difficult-to-treat depression: differential diagnosis, subtypes, and comorbidities. *Journal of Clinical Psychiatry* 70 Suppl 6, 10-15.
- Goekoop, J.G., Hoeksema, T., Knoppert-van der Klein, E.A., Klinkhamer, R.A., van Gaalen, H.A., Van Londen, L., De Weme, R., Zwinderman, A.H., 1992. Multidimensional ordering of psychopathology. A factor-analytic study using the Comprehensive Psychopathological Rating Scale. *Acta Psychiatrica Scandinavica* 86, 306-312.
- Grilo, C.M., Stout, R.L., Markowitz, J.C., Sanislow, C.A., Ansell, E.B., Skodol, A.E., Bender, D.S., Pinto, A., Shea, M.T., Yen, S., Gunderson, J.G., Morey, L.C., Hopwood, C.J., McGlashan, T.H., 2010. Personality disorders predict relapse after remission from an episode of major depressive disorder: a 6-year prospective study. *Journal of Clinical Psychiatry* 71, 1629-1635.
- Guy, W., 1976. *ECDEU Assessment Manual for Psychopharmacology*. National Institute of Mental Health, Rockville, MD.
- Kadouri, A., Corruble, E., Falissard, B., 2007. The improved Clinical Global Impression Scale (iCGI): development and validation in depression. *BMC Psychiatry* 7, 7.
- 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.
- Katon, W.J., Lin, E.H., Von, K.M., Ciechanowski, P., Ludman, E.J., Young, B., Peterson, D., Rutter, C.M., McGregor, M., McCulloch, D., 2010. Collaborative care for patients with depression and chronic illnesses. *New England Journal of Medicine* 363, 2611-2620.
- Kessler, R.C., Berglund, P., Demler, O., Jin, R., Merikangas, K.R., Walters, E.E., 2005. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry* 62, 593-602.

- Klinkman, M.S., 2009. Assessing functional outcomes in clinical practice. *American Journal of Managed Care* 15 Suppl 11, S335-S342.
- Lecrubier, Y., 1997. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *European psychiatry* 12, 224-231.
- Leiknes, K.A., Finset, A., Moum, T., Sandanger, I., 2007. Course and predictors of medically unexplained pain symptoms in the general population. *Journal of Psychosomatic Research* 62, 119-128.
- van der Lem, R., van der Wee, N.J., van Veen, T., Zitman, F.G., 2011. The generalizability of antidepressant efficacy trials to routine psychiatric out-patient practice. *Psychological Medicine* 41, 1353-1363.
- Luanaigh, C.O., Lawlor, B.A., 2008. Loneliness and the health of older people. *International Journal of Geriatric Psychiatry* 23, 1213-1221.
- Maier, W., Falkai, P., 1999. The epidemiology of comorbidity between depression, anxiety disorders and somatic diseases. *International Clinical Psychopharmacology* 14 Suppl 2, S1-S6.
- Massion, A.O., Dyck, I.R., Shea, M.T., Phillips, K.A., Warshaw, M.G., Keller, M.B., 2002. Personality disorders and time to remission in generalized anxiety disorder, social phobia, and panic disorder. *Archives of General Psychiatry* 59, 434-440.
- McKenzie, M., Clarke, D.M., McKenzie, D.P., Smith, G.C., 2010. Which factors predict the persistence of DSM-IV depression, anxiety, and somatoform disorders in the medically ill three months post hospital discharge? *Journal of Psychosomatic Research* 68, 21-28.
- Melartin, T.K., Rytala, H.J., Leskela, U.S., Lestela-Mielonen, P.S., Sokero, T.P., Isometsa, E.T., 2004. Severity and comorbidity predict episode duration and recurrence of DSM-IV major depressive disorder. *Journal of Clinical Psychiatry* 65, 810-819.
- Meyers, B.S., Sirey, J.A., Bruce, M., Hamilton, M., Raue, P., Friedman, S.J., Rickey, C., Kakuma, T., Carroll, M.K., Kiosses, D., Alexopoulos, G., 2002. Predictors of early recovery from major depression among persons admitted to community-based clinics: an observational study. *Archives of General Psychiatry* 59, 729-735.
- Mezuk, B., Eaton, W.W., Albrecht, S., Golden, S.H., 2008. Depression and type 2 diabetes over the lifespan: a meta-analysis. *Diabetes Care* 31, 2383-2390.
- Montgomery, S.A., Åsberg, M., 1979. A new depression scale designed to be sensitive to change. *The British Journal of Psychiatry* 134, 382-389.
- van Noorden, M.S., Giltay, E.J., den Hollander-Gijsman, M.E., van der Wee, N.J., 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.
- Penninx, B.W., Nolen, W.A., Lamers, F., Zitman, F.G., Smit, J.H., Spinhoven, P., Cuijpers, P., de Jong, P.J., van Marwijk, H.W., van der Meer, K., Verhaak, P., Laurant, M.G., de Graaf, R., Hoogendijk, W.J., van der Wee, N.J., Ormel, J., van Dyck, R., 2011. Two-year course of depressive and anxiety disorders: Results from the Netherlands Study of Depression and Anxiety (NESDA). *Journal of Affective Disorders* (in press).
- Phillips, K.A., Pagano, M.E., Menard, W., Fay, C., Stout, R.L., 2005. Predictors of remission from body dysmorphic disorder: a prospective study. *Journal of Nervous and Mental Disorders* 193, 564-567.
- Robins, C.J., Chapman, A.L., 2004. Dialectical behavior therapy: current status, recent developments, and future directions. *Journal of Personality Disorders* 18, 73-89.

- Roy-Byrne, P., Craske, M.G., Sullivan, G., Rose, R.D., Edlund, M.J., Lang, A.J., Bystritsky, A., Welch, S.S., Chavira, D.A., Golinelli, D., Campbell-Sills, L., Sherbourne, C.D., 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.
- Rush, A.J., Trivedi, M.H., Wisniewski, S.R., Nierenberg, A.A., Stewart, J.W., Warden, D., Niederehe, G., Thase, M.E., Lavori, P.W., Lebowitz, B.D., McGrath, P.J., Rosenbaum, J.F., Sackeim, H.A., Kupfer, D.J., Luther, J., Fava, M., 2006. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *The American Journal of Psychiatry* 163, 1905-1917.
- Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., 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.
- Skodol, A.E., Grilo, C.M., Keyes, K.M., Geier, T., Grant, B.F., Hasin, D.S., 2011. Relationship of Personality Disorders to the Course of Major Depressive Disorder in a Nationally Representative Sample. *The American Journal of Psychiatry* 168, 257-264.
- Souery, D., Oswald, P., Massat, I., Bailer, U., Bollen, J., Demyttenaere, K., Kasper, S., Lecrubier, Y., Montgomery, S., Serretti, A., Zohar, J., Mendlewicz, J., 2007. Clinical factors associated with treatment resistance in major depressive disorder: results from a European multicenter study. *Journal of Clinical Psychiatry* 68, 1062-1070.
- Speckens, A.E., van Hemert, A.M., Bolk, J.H., Rooijmans, H.G., Hengeveld, M.W., 1996. Unexplained physical symptoms: outcome, utilization of medical care and associated factors. *Psychological Medicine* 26, 745-752.
- Trivedi, M.H., 2001. Sensitizing clinicians and patients to the social and functional aspects of remission. *Journal of Clinical Psychiatry* 62 Suppl. 19, 32-35.
- Vatnaland, T., Vatnaland, J., Friis, S., Opjordsmoen, S., 2007. Are GAF scores reliable in routine clinical use? *Acta Psychiatrica Scandinavica* 115, 326-330.
- de Waal, M.W., Arnold, I.A., Eekhof, J.A., van Hemert, A.M., 2004. Somatoform disorders in general practice: prevalence, functional impairment and comorbidity with anxiety and depressive disorders. *The British Journal of Psychiatry* 184, 470-476.
- Ware, J.E., Sherbourne, C.D., 1992. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care* 30, 473-483.
- Watson, D., Weber, K., Assenheimer, J.S., Clark, L.A., Strauss, M.E., McCormick, R.A., 1995. Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. *Journal of Abnormal Psychology* 104, 3-14.
- Weinberger, M.I., Sirey, J.A., Bruce, M.L., Heo, M., Papademetriou, E., Meyers, B.S., 2008. Predictors of major depression six months after admission for outpatient treatment. *Psychiatric Services*, 59, 1211-1215.
- Wittchen, H.U., Jacobi, F., 2005. Size and burden of mental disorders in Europe--a critical review and appraisal of 27 studies. *European Neuropsychopharmacology*, 15, 357-376.
- Yonkers, K.A., Dyck, I.R., Warshaw, M., Keller, M.B., 2000. Factors predicting the clinical course of generalised anxiety disorder. *The British Journal of Psychiatry*, 176, 544-549.

Chapter 7

Summary, discussion and future perspectives

Summary

This thesis focused on *real-world* patients and *real-world* outcomes by using Routine Outcome Monitoring (ROM) in patients with Mood, Anxiety, and Somatoform (MAS) disorders. The primary aims of the series of studies were to investigate correlates of disease characteristics in a large cohort of treatment-seeking patients with MAS disorders, and to discuss these findings in relation to existing data derived from RCTs and general population studies. In addition, using prospective data, we investigated whether baseline characteristics measured with ROM could predict outcomes in real-world or naturalistic treatment settings. The secondary aim was to assess the feasibility of conducting large-scale clinical epidemiological research with ROM data that had been collected in everyday clinical practice, thereby representing the phenotype of ‘real’ patients. In the remainder of this chapter we will summarise the major findings of the presented studies, discuss these findings and ROM in a broader perspective, and conclude with thoughts and recommendations about future studies.

Summary of major findings

The first study of this thesis (Chapter 2) concentrated on gender differences in patients with Major Depressive Disorder (MDD). Although gender differences in course and symptomatology of depression have frequently been studied, a closer look at the literature reveals that almost all studies have been conducted in population samples or clinical trial samples. In a recent study of our group we have shown that only 20-25% of our MDD patients would qualify for inclusion in a regular clinical trial (van der Lem et al., 2010), so we argued that patients from clinical trials do not adequately represent treatment-seeking everyday or real-world patients. In other words, external validity is limited. The aim of this study therefore was to investigate gender differences in clinical characteristics of a naturalistic sample of MDD patients from our ROM database. We used ROM baseline data of 1,131 MDD patients, and analysed symptom severity, symptom profiles, comorbidity and general health status. The most remarkable findings were that women reported a more severe ‘subjective’ phenotype on self-rated scales, with more symptoms and more severe symptoms of depression and a lower general health status than men. On more objective observer-rated scales, however, no significant differences in symptom severity were found between women and men. These findings suggest that the burden of MDD is higher in women than in men, but only subjectively. The pattern of subjective symptomatology

was also different in women compared to men, with women reporting more loss of energy, hypersomnia, feelings of worthlessness, loss of appetite, fatigue, loss of interest in sex, feelings of guilt and crying. The majority of these findings were largely consistent with previous studies in non-naturalistic patient samples. In Chapter 3, we extended the cross-sectional analyses on ROM baseline assessments in patients with MDD. In this study we compared clinical characteristics of patients with a pre-adult onset versus adult-onset MDD. Evidence from earlier studies pointed towards a more severe phenotype in pre-adult onset MDD. However, these findings again were mostly derived from studies in either population-based studies or clinical trials. In multivariable analyses in 1,105 MDD patients, we found that patients with pre-adult onset MDD more often had a history of suicide attempts and current suicidal thoughts compared with adult-onset MDD patients. This important study finding replicated earlier findings in clinical samples and supports the hypothesis that pre-adult MDD is characterised by a distinct phenotype compared with adult-onset MDD. Other previous findings, e.g. having more severe symptoms (Zisook et al., 2004; Zisook et al., 2007) or anxiety (Parker et al., 2003) in pre-adult onset MDD, could not be replicated. This is important to ascertain, since our samples more truly represent 'real' patients than previous studies did. Although possibly confounded by recall bias and measurement error, the age of onset of MDD can be obtained relatively easy and may help the clinician in risk-assessment of suicidality.

After these first two studies in MDD patients we expanded our cross-sectional analyses to all patients with MAS disorders. In Chapter 4 we studied prevalence and correlates of deliberate self-harm and suicidal ideation (DSHI) in a large group of patients referred for treatment of either a mood, anxiety or somatoform disorder, or a combination of these diagnoses. We analysed 2,844 patients with a ROM baseline assessment and used the self-harm subscale of the Dimensional Assessment of Personality Pathology-Short Form (DAPP-SF) to assess DSHI. Lifetime DSHI was reported by 55% of the total sample. Being married, having a higher educational status and having high anxiety symptoms were independently associated with lower risks of DSHI. Having a higher number of MAS diagnoses, having more depressive symptoms and higher scores on the emotional dysregulation subscale of the DAPP-SF were associated with higher risks of DSHI. The negative association with anxiety measured with the Brief Anxiety Scale (BAS) in adjusted analyses was unexpected. The most plausible explanation of this finding would be that patients with high anxiety levels but similar levels of depressive symptoms were too afraid to hurt themselves. These results in a unique broad sample of outpatients with common mental disorders demonstrate that DSHI is a common phenomenon in MAS outpatients, with distinct clinical correlates. The findings may help clinicians to identify patients at

increased risk for DSHI and completed suicide.

In the next part of this thesis we shifted the focus to clinical predictors of real-world outcomes, using prospective ROM measurements in addition to the baseline assessments. In Chapter 5 we used ROM baseline and follow-up assessments of up to 24 months to investigate whether individual baseline depressive symptoms would predict outcome in 1,489 MDD patients. In earlier studies, many clinical correlates of MDD-treatment outcome have been identified. Earlier age of onset, more severe complaints and cluster B and C personality traits were among the predictors of less favourable outcome (see for example Frank et al., 2010; Penninx et al., 2011; Souery et al., 2007; Vuorilehto et al., 2009). Most studies had a limited follow-up time of three months to one year. Moreover, the predictive value of specific depressive symptoms at baseline had not been systematically assessed before. In daily clinical practice, being able to predict which patients are at risk to become non-responders or develop a chronic course of the depression is highly important. We analysed the predictive value of all 21 individual depressive symptoms measured with the widely used Beck depression Inventory-Revised (BDI-II; Beck et al., 1988) on naturalistic treatment outcome, defined as remission (score <10) or response ($\geq 50\%$ reduction) on the Montgomery-Åsberg Depression Rating Scale (MADRS; Montgomery, 1979). Using multivariable Cox regression analyses, we found that after adjusting for known clinical variables correlating with treatment outcome, baseline presence of the two major symptoms 'pessimism' and 'loss of energy' strongly and independently predicted a higher chance of a chronic course. In daily clinical practice, the presence or absence of these symptoms are relatively easy to assess, and the presence of pessimism and/or loss of energy should alert the clinician to the higher risk of a chronic MDD course.

In the final study (Chapter 6), we extended the prediction of real-world outcomes to the total group of patients with MAS disorders. In the DSM-IV-TR, MAS disorders show a considerable overlap in diagnostic criteria and in daily practice mutual comorbidity of these disorders is substantial. Although in most clinical trials only patients with certain primary axis-I diagnoses are included, the concept of primary diagnosis is rather vague (van der Lem et al., 2010). The frequent occurrence of MAS-comorbidity complicates the establishment of the primary or main diagnosis. Hence, in daily clinical practice, a patient with for example a MDD and a social phobia is treated according to the MDD guideline, the social phobia guideline, or a combination of both. In this particular study we aimed to assess predictors of 'cross-diagnostic' outcome in clinical practice. We defined outcome as a certain reduction on a combination of two disorder independent measurement scales: the observer-rated Clinical Global Improvement Scale (CGI; Kadouri et al., 2007) and the self-report Brief Symptom Inventory (BSI; Derogatis and Melisaratos, 1983).

We used baseline ROM data as well as prospective data with up to 24 months of follow-up. We used a sample of 892 MAS patients, and validated our results in an independent sample of 1,392 MAS patients. In multivariable Cox regression analyses, we found that after adjusting for age and gender, the following characteristics were independent predictors of unfavourable outcome: advancing age, having MAS comorbidity or a somatoform disorder, the personality traits intimacy problems, affect lability and self-harm, and a low self-reported general health status. These findings were confirmed in a second sample of MAS patients, and could help the clinician to identify patients who are at risk for a chronic course of the MAS disorder. Whether these patients, based on the above profile, could benefit from extra interventions should be the focus of future studies.

General discussion

Routine Outcome Monitoring is the systematic assessment of treatment outcomes in everyday clinical practice. In psychiatric specialty care, ROM assessments include standardised and validated measurement instruments. The ROM-infrastructure in Leiden uses a broad test battery that comprises a range of both self-report and observer-rated scales. The test battery includes a diagnostic instrument, instruments that measure generic and disorder-specific symptom severity, and scales that measure psychosocial functioning. In general, ROM is primarily used as a tool for patient and clinician to monitor treatment progress. Although more specific research is necessary, studies indicate a positive impact of ROM and feedback on mental health status (Carlier et al., 2010). In anonymised form, the systematically obtained data can also be used for clinical epidemiological research purposes (de Beurs et al., 2011), as exemplified by this thesis. Whenever a broad range of instruments is applied, detailed phenotypic information of patients can be derived from ROM. In the Leiden region in the Netherlands, ROM assessments are part of the routine diagnostic and therapeutic procedure in patients with MAS disorders in the outpatient clinics at the Leiden University Medical Center (LUMC) and Rivierduinen (RD; see Box 1.1, Chapter 1). In Leiden, assessments ideally take place directly after referral (ROM baseline assessment) and at any step or change in the treatment protocol. In practice, this is on average every 3-4 months repeatedly during treatment. In Leiden, only patients who are unable to undergo ROM assessments due to language problems, disease severity or refusal are excluded from ROM. With approximately 80% of all patients taking part in the Leiden ROM, results are highly generalisable to everyday clinical practice.

The Leiden ROM infrastructure is an example of an extensive ROM infrastructure. This extensive ROM has several advantages, e.g. detailed phenotyping, staging of psychiatric disease, identification of residual symptoms ('rollback phenomena', see Fava et al. 2011). Furthermore, it provides opportunities for clinical epidemiological research and benchmarking. However, important differences exist among implemented ROM systems worldwide. For example, some institutions apply routine monitoring with the use of only one self-report instrument, that is administered at every visit (e.g. Zimmerman et al., 2005).

The studies presented in this thesis focused on clinical characteristics and predictors of outcome in everyday patients with mood, anxiety and somatoform disorders. Although this seems a trivial fact, studies in naturalistic samples of MAS patients in psychiatric specialty care were scarce. The large-scale ROM infrastructure in the Leiden region provided a unique opportunity to investigate important clinical aspects in a relatively unselected treatment-seeking outpatient population, yielding findings that are easy to generalise to routine clinical outpatient practice. This contrasts with findings from clinical trials that included highly selected patient samples resulting in a more limited generalisability. Findings from population studies do neither reflect a treatment-seeking population. In fact, 75-80% of our ROM MDD-patients would not have met the in- and exclusion criteria for most clinical trials (van der Lem et al., 2010).

The findings of the explorative studies in this thesis have provided unique insights in the phenomenology of 'real' MAS patients in daily clinical practice. Predictors of naturalistic outcome have been established. In addition, the strengths and weaknesses of a large-scale and extended ROM initiative have become tangible. In the next paragraphs, we will discuss our findings in a broader perspective.

Phenotype of real-world MAS patients

In the process of designing the studies presented in this thesis, we decided to start our explorative analyses with a well-studied topic in psychiatric literature. As mentioned before, gender differences in depression have been studied extensively. The naturalistic character of our study sample was a new point of view. The results of this first study partly confirmed and extended previous research. The most important implication of the first study was that for a balanced ROM assessment both observer-rated and self-report scales are necessary, which supported earlier work of Möller (2009). If the ROM assessments would have comprised only self-rating instruments, important nuances would have been missed. After all, we found no differences in number nor severity of depressive symptoms

between men and women on observer-rated scales, while substantial differences existed on self-rated scales. This stresses the need for maintaining the observer-rated scales in ROM, even in an era of inevitable cost reductions in healthcare. Another important conclusion of the first study in this thesis was that the quality of the anonymised patient data provided by our ROM database was successfully applied for the intended analyses.

The main finding of the second study was that MDD patients with disease onset before the age of 18 years were more suicidal compared to those with onset at adult age. This was in part a replication of earlier studies in different samples. The fact that in multivariable analyses all differences in clinical profile, except suicidality, became non-significant could at least in part explain the high variability of the findings in earlier studies, which not always had included a comprehensive set of potential confounders. The higher rates of current and past suicidality in pre-adult onset MDD may have important clinical implications, as completed suicide needs to be prevented in patients with psychiatric disorders, and it is one of the leading causes of death worldwide. The estimated global burden of suicide is a million deaths per year (WHO, 2002), or 14.5 deaths per 100.000 people worldwide. These are probably underestimations. Large differences exist between countries, and men are much more at risk for completed suicide with a male-female ratio of 2-4 to 1 in developed countries (WHO, 2002). Previous self-harm or suicide attempts are major risk factors for completed suicide (Cavanagh et al., 2003; Coryell and Young, 2005). Based on psychological autopsies in completed suicide patients, it is estimated that 90% (Cavanagh et al., 2003) or even more (Ernst et al., 2004) had a DSM-IV Axis I psychiatric disorder at the time of suicide. More than half of the patients who die of suicide meet criteria for current MDD (Cavanagh et al., 2003; Hawton and van Heeringen, 2009), and about 4% of MDD patients die by suicide. Despite dramatic increases in awareness, prevention and treatment efforts, no significant decrease in suicide rates has occurred in the US from 1990 to 2003 (Kessler et al., 2005). These figures stress the need for early and better identification of suicidality, especially in MDD patients, and the need for specific interventions. Based on the results of Chapter 2, clinicians should be especially aware of suicidality in male MDD patients with pre-adult onset MDD and previous suicide attempts.

The finding of more suicidality in pre-adult onset MDD patients increased our interest in the phenomenon of suicidality and the broader concept of deliberate self-harm. Since suicidal patients and patients who engage in self-destructive behaviour are excluded from most clinical trials in psychiatry, we investigated these phenomena in depth in our total group of MAS patients. This third study also in part replicated earlier studies in more strictly selected populations, but also yielded new insights. The most important outcome was the apparent protective effect of state anxiety on deliberate self-harm and

suicidal ideation. Of course, this finding needs replication in future studies. Furthermore, the broad sample of 2,844 treatment-seeking MAS patients with substantial comorbidity allowed us to identify correlates of self-harm irrespective of DSM-IV diagnosis, which increased applicability in clinical practice.

Predictors of real-world outcomes

One of the most important topics for both patients and clinicians regards the chances of getting better. This is of course not specific for psychiatric patients. Large-scale epidemiological studies have identified risk factors for adverse outcomes in a wide range of somatic diseases or disease variants. These studies have resulted in risk-assessment schedules that have been widely adapted in worldwide guidelines. A famous example is the Framingham heart study, a cohort study of over 5,000 inhabitants of Framingham, Massachusetts. This study, initiated in 1948, has calculated detailed chances of cardiovascular events if a patient corresponds to a certain profile, based on gender, age, systolic blood pressure, and serum lipids (Anderson et al., 1991; Dawber et al., 1951). Based on this study, with numerous publications and an offspring study (Kannel et al., 1979), patients with certain risk profiles are considered as subgroups for which risk scores can be calculated and different treatments can be advised. International guidelines have adapted these risk calculators. In Europe, the Systematic Coronary Risk evaluation (SCORE) project has used data of over 200.000 European patients to calculate estimations of ten-year risk of fatal cardiovascular disease (Conroy et al., 2003). As discussed in the introduction of this thesis, for psychiatric disorders, those detailed studies have never been performed. Based on the available epidemiological literature, only rough estimations about the course of the disease can be provided for the total group of patients with a certain disorder (see for example Judd, 1997). Until now, guidelines and treatment algorithms only roughly differentiate between subgroups of patients. For example: The Dutch Multidisciplinary Guideline for the treatment of MDD (www.ggzrichtlijnen.nl) discriminates between first episode or recurrent episode, and between mild or moderate/severe depression. Furthermore, patients with moderately to severe depression have the options of a modern antidepressant or psychotherapy, whereas the treatment algorithm for patients with Major Depression with psychotic features suggests treatment with a tricyclic antidepressant, with possible addition of an antipsychotic agent. Depression with atypical features may better respond to monoamine-oxidase inhibitors than other classes of antidepressants. However, also for psychiatrists, being able to predict the chance of remission in more

sophisticated subgroups of patients is of utmost importance, yet still very imprecise. Based on the findings of the studies in this thesis, in the next part of this discussion we present a theoretical future algorithm of risk assessment in MDD (figure 7.1).

The fourth study of this thesis can be interpreted as an attempt to predict outcomes of MDD in a real-world treatment setting based on baseline symptoms that can easily be obtained. We have demonstrated that ROM assessments can be helpful in risk assessment and that the baseline presence of both lack of energy and pessimism symptoms measured with the widely used BDI-II, were highly predictive of non-remission of MDD after a follow-up of up to 2 years. In the fifth and final study of this thesis we aimed to establish predictors of poor outcome in the whole group of patients with MAS disorders. The broad phenotypic assessment with ROM allowed us to closely investigate demographical and clinical risk factors of poor outcome in a naturalistic treatment setting. Furthermore, we used a cross-diagnostic design in which we included all MAS patients thereby abandoning the rather vague concept of primary diagnosis often used in studies. This made the results of this study easy and widely applicable to clinical practice. Although much more research is needed, our longitudinal ROM studies can be seen as one further step towards a more detailed risk assessment, based on extensive phenotypic data. In the future, hopefully, adding genotypes and endophenotypes (biological substrates) to phenotypic data may further enhance the precision of risk-assessment. In addition, important environmental information that might be important in epigenetic effects (e.g. childhood trauma, daily stress) may also be ascertained with ROM. Efforts in finding biological substrates of symptom dimensions of depression and anxiety have already been published (Veen et al., 2011; Wardenaar et al., 2011).

ROM and Comparative Effectiveness Research

In the literature, research based on ROM data is often regarded as ‘patient-centered research’. If treatment details are taken into account, ROM-data driven research could be used for Comparative Effectiveness Research (CER). CER is designed to improve the clinical decision-making process by providing research evidence on the effectiveness and risk-benefit profile on different therapeutic options for specific patient subpopulations (Mane et al., 2011; Sox and Greenfield, 2009). In the US, the Patient-Centered Outcomes Research Institute (PCORI) has been established to facilitate CER. The mission of this institute is to help people make informed healthcare decisions, and to improve healthcare delivery and outcomes (Washington and Lipstein, 2011). The US government has offered

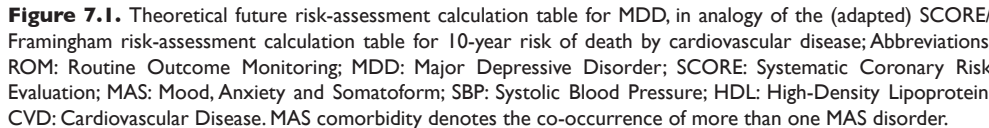
a funding of more than 3 billion US Dollar for the next decade, allowing the support of numerous studies of CER, in different areas of healthcare. In Europe, to our knowledge, no comparable large-scale initiatives are being developed. In theory, ROM databases of different institutions could be merged in large collaboration efforts and used for CER. For such overarching initiatives to be successful, many consensus steps have still to be taken. On a national level, for The Leiden Institute for the Advancement and Integration of ROM (LIAIROM), achieving consensus about implementation of ROM and e.g. choice of measurement instruments in order to integrate ROM databases is a priority area. For example: based on our studies, specific interventions for MDD patients meeting the profile of advanced age, comorbidity, cluster B traits and poor reported general health could be investigated. Or if treatment details become available in ROM, it would be possible to investigate which antidepressant is associated with the highest chance of remission in patients with baseline pessimism and lack of energy.

A theoretical future algorithm of risk assessment and treatment recommendations in MDD patients which resembles the cardiovascular disease risk calculation is presented in figure 7.1.

In cardiovascular risk management differentiated treatment recommendations (e.g. lifestyle adaptations with or without treatment with statins) are based on several phenotypic and endophenotypic aspects (e.g. smoking and systolic blood pressure). One could imagine that in future psychiatric practice, comparable treatment recommendations would be possible (e.g. lifestyle adaptations, first-step interventions, versus psychotherapy or pharmacotherapy or extra interventions), based on ROM assessments comprising detailed phenotypic information and information about environmental factors, combined with genotypes and endophenotypes (e.g. cortisol levels).

ROM critically appraised

One of the aims of this thesis was to assess the feasibility of conducting large-scale clinical epidemiological research with data provided by ROM. The studies presented here demonstrate that the ROM infrastructure in RD/LUMC provide vast possibilities in conducting both cross-sectional and prospective analyses, with the opportunity of studying specific phenotypes of MAS patients and risk factors of outcomes and disease course. Several aspects deserve further notification. We will now focus on the strengths and weaknesses of ROM. We will differentiate between the Leiden ROM system and ROM in general.



The Leiden ROM infrastructure has been developed and implemented with the combined goals of treatment monitoring for patient and clinician, and to serve as a large database for research and for benchmarking. The use of computerised assessments and dedicated software prevents missing data within specific instruments, and provides high data quality. Another strength is the fact that assessments are performed by specially trained research nurses who are not involved in treatment, ensuring neutral assessments without the risk of measurement bias that could occur when clinicians would rate their own patients. Moreover, the extensive baseline assessments, including a comprehensive and well validated diagnostic measurement instrument (MINI-Plus), an instrument that assesses dysfunctional personality traits (DAPP-SF) and repeatedly administered scales that cover the symptom level and psychosocial functioning, provide detailed phenotypic information at every phase of treatment. Both phenotypic research on a group level and outcomes research can be performed. The choice of measurement instruments at the time of implementation (from 2002) was based on psychometric quality and applicability of the instrument, and on the availability in the public domain. For longitudinal research, obviously, the test battery should remain extensive and ideally unchanged. Unfortunately, in the past years copyright issues forced us to abandon certain scales (e.g. DAPP-SF, BSI). The high rate of ROM baseline assessments of approximately 80% is another strength. Because ROM is integrated in everyday clinical practice, only patients who are not fluent in Dutch or those patients who are too ill to undergo assessments are excluded from ROM.

By utilising ROM data for the studies in this thesis we encountered several limitations of the Leiden ROM infrastructure. The most important limitation of the present Leiden ROM system is the current lack of integration with the (anonymous) electronic patient record forms (PRFs), preventing the researcher from retrieving information about exact treatment modality, detailed psychiatric history, family history of psychiatric diseases, and somatic status. In order to obtain detailed information about e.g. treatment, the researcher has to hand-search the patient record files. This hand-searching could compromise anonymity. Integration of ROM with PRFs would have given us the opportunity to adjust for treatment modality or guideline adherence in the two longitudinal studies (Chapters 5 and 6). In these studies we have not been able to adjust for possible confounding by the treatment type, patient characteristics, and disease characteristics. On the patient level, integration of ROM and PRF could improve decision making by identifying subpopulations of patients with similar characteristics and to determine which treatments have been successful with minimal side-effects (Iglehart, 2009; Mane et al., 2011). This type of research is commonly regarded as Comparative Effectiveness Research (CER). Another limitation of our present ROM system regards the time intervals of ROM assessments. From a clinical point of view,

ideally follow-up assessments should be scheduled before every next treatment step. From a research perspective, ROM measurements at fixed time intervals, independent from disease severity are to be preferred for most longitudinal data analytic techniques. Cox regression analysis, the method we used in Chapter 5 and 6 for predicting outcome, is an example of a technique that is able to use variable numbers of measurements and variable time-intervals. In the current clinical practice in Leiden, measurements are initiated by the research nurse, the clinician or patient, but not always before a change of treatment or a next treatment step according to the guidelines. This unpredictable interval of ROM assessments without automatically scheduled re-assessments may contribute to bias due to attrition and complicates data analysis and interpretation. At present, efforts are made to incorporate a more systematic measurement interval in the Leiden ROM.

A potential limitation or pitfall of ROM in general is important to consider. In this modern era of excessive growth of healthcare and inevitable health-costs, the power of health insurance companies is growing and the professional autonomy of medical specialists is increasingly under pressure. Understandably, policy makers like health insurance companies and governments demand more and more insight in costs of treatments and treatment processes to be able to control these costs. Since ROM is a potentially valuable source of information regarding these processes, many health providers have implemented ROM initiatives over the past years, with benchmarking as one of the major goals. The studies in this thesis may illustrate that benchmarking based on ROM assessments is possible but could also be a hazardous operation. First of all, some institutions have adapted a ROM system in which only a succinct set of outcome scales or only one scale is used. Even with our extended ROM measurement scales, it would be hard to derive reliable benchmarking data because of important inter-patient differences that require complex statistics to take into account. For example, if in a certain clinic more MDD patients with comorbid personality disorders are being treated, outcomes may be worse compared to another clinic that uses the same guidelines but where patients with less complicated complaints seek treatment. Our major concern would be that in the case of a limited ROM assessment battery, policy makers will draw conclusions based on insufficient or inadequately analysed data. For example, in tertiary care clinics or specialised secondary care clinics, typically patients with treatment-resistant complaints, somatic comorbidity, co-existent personality pathology, or a combination of these are being treated. Those patients are likely to have worse treatment-outcomes, irrespective of the quality of treatment, as compared with less complicated patients in general psychiatric specialty care.

In this thesis we have shown that for a balanced ROM assessment it is necessary to apply both self-report and observer-rated scales. Ironically, due to the aforementioned efforts to control healthcare costs, the current tendency is to limit or even omit the observer-rated scales and standardised diagnostic interviews from ROM. To base clinical and benchmark decisions solely on self-report scales would perhaps save salary costs of research nurses. However, eventually these self-report measurements could only reflect part of the process of diagnostics and treatment outcome because more objective information derived from observer-rated scales would be lacking. In this thesis, we have shown that ROM, including observer-rated measurements, provides valuable and more objective information. Finally, the large variety of ROM assessment scales that are being used by different health organisations limits comparability and is another caveat with respect to benchmarking. Calculating standardized scores or z-scores could help with respect to this problem.

A last limitation of ROM in general to discuss here relates to the population under study. The studies described here focused on patients with MAS disorders, also regarded as common mental disorders. Our local decision to implement ROM firstly in MAS disorder patients was based on the research profile of our department. However, substantial overlap of symptoms and mutual comorbidity of MAS disorders, the wide availability of validated rating scales in the MAS domain, and the fact that MAS patients were expected to be more compliant than e.g. psychotic patients or patients with personality disorders as main diagnosis, also played a role in this decision. Indeed, in several ROM initiatives in other patient populations, e.g. psychotic patients or Severe Mental Illness (SMI)-patients, these considerations have been underlined (Mulder et al., 2010).

Future Perspectives

The studies presented in this thesis can be regarded as a first set of explorative studies to demonstrate the possibilities of conducting clinical epidemiological research with ROM data. Several adjustments, e.g. adding treatment details in ROM, may further improve the potential of future studies in our setting (see table 7.1).

Besides for clinical epidemiological research, the Leiden ROM data can serve as basis for research in other domains. Examples of these are biological and psychometric research. In Leiden, many researchers already use the ROM-database for studies in these domains. Results have been published and will be in the near future. To further illustrate the potential of ROM, we will give some examples of current projects that use the Leiden ROM infrastructure.

First, the lack of well-known biological markers in psychiatric disorders complicates the borders of disease. When can someone with certain complaints be classified as a patient? The diagnostic classification system DSM-IV only partially answers that question by operationalising disorders by consensus definitions. Due to the absence of clear markers and borders, the line between 'healthy' and 'sick' will be hard to define. Validated measurement scales are helpful in defining and establishing that line. However, for many validated measurement scales used in ROM no reference values in the general population have been calculated. NormQuest is a study initiated in Leiden that aims to assess those reference values for the commonly used measurement scales in MAS patients (Schulte et al., in press).

Second, the ROM infrastructure with naturalistically obtained data also allows for add-on research. The Mood, Anxiety, Somatoform disorders and Hypothalamus pituitary adrenal (HPA) axis Biobank (MASHBANK) is an example of this type of research. The MASHBANK has been founded to investigate the link between genetic variants, functioning of the HPA axis and the phenotype in patients with MAS disorders, and comparing those with patients from the general population. Patients routinely enrolled in ROM have been asked informed consent to donate DNA for the MASHBANK, after MEC approval of the protocol. So far, almost 2000 samples of MAS patients and control subjects have been collected.

Third, as discussed earlier, ROM measurement scales should ideally be free of copyright in order to allow broad implementation. For longitudinal research, a stable test battery is paramount. Unfortunately, the past years publishers have claimed copyrights for measurement scales used in our ROM test battery forcing us to abandon those scales. The ROM Research Center (COROM) in Leiden facilitates ROM-related research of which the development of new measurement scales for the free domain is one of the topics. The development and validation of freely available scales comparable to copyright-protected scales are needed. An example of the latter is the construction of a new measure for the assessment of psychological distress. We developed and validated the Symptom Questionnaire-48 (SQ-48), a measurement scale that assesses 48 symptoms in nine symptom domains (Carlier et al., in preparation). Furthermore, translation and subsequent validation of measurement scales in different languages is necessary to allow non-Dutch patients to be enrolled in ROM.

As mentioned before, the studies in this thesis are examples of the potential of ROM. Many more studies and collaborations are necessary to develop useful risk calculators, to integrate biological markers in ROM, and to develop specific interventions for subgroups of patients. Although practical obstacles may have to be faced, lessons

from the cardiovascular field (Framingham, SCORE) but also oncology demonstrate that large-scale collaboration may dramatically improve outcomes step by step in large groups of patients. For example, acute leukemia is the most common form of childhood cancer, comprising approximately 30% of all malignancies in children. Survival rates for Acute Lymphatic Leukemia have increased dramatically since the 1980s, with current five-year overall survival rates of over 85% (Gatta et al., 2005; Pui et al., 2004; Pui and Evans, 2006). These improved survival rates are due to large-scale collaborations and treatment of large groups of patients according to standardised research protocols, based on staging of disease and constant monitoring of outcomes. These protocols have evolved over and over according to outcomes of trials and findings of more biological studies (Lee et al., 2000; Pui and Evans, 2006). In psychiatry, establishment of the international schizophrenia consortium (ISC) has resulted in large-scale genetic studies in schizophrenia and bipolar disorder (e.g. Purcell et al., 2009). In Europe, the GENDEP consortium aims to use genetic profiles to predict outcome of antidepressant treatment (e.g. Uher, et al., 2010). These initiatives demonstrate that large-scale collaborations are possible. Indeed, initiatives for staging and profiling in psychiatry are emerging. In a recent review of Fava et al. (2011), the authors stress the importance of staging in psychiatry. "Staging differs from conventional diagnostic practice in that it not only defines the extent of progression of a disorder in a particular point in time but also reveals a person's current location on the continuum of the course of illness". Detailed information about symptoms, comorbidity, psychosocial functioning, quality of life and response to treatment derived by ROM assessments could serve as basis for staging.

Concluding remarks

The studies in this thesis demonstrate that a proper ROM infrastructure is a valuable source for clinical epidemiological research in naturalistic or 'real' patients. We have studied in detail aspects of the phenotype of MAS patients, and established risk factors of poor outcome in both MDD and a treatment-seeking MAS population. Major findings were the phenotype of more suicidality in pre-adult onset MDD patients, and correlates of deliberate self-harm in a large group of MAS patients. Pessimism and lack of energy were found to be predictors of poor outcome in MDD patients, and advancing age, dysfunctional cluster-B personality traits, MAS comorbidity and poor reported general health status were risk factors for poor outcome in MAS disorders. These risk factors need future replication in other naturalistic cohorts, and could result in detailed risk-assessment

calculators comparable to those used in cardiovascular medicine. Integration of treatment details in ROM would allow for comparative effectiveness studies in specific subgroups of patients. Integration of biological markers in ROM and (inter)national collaboration could bring ROM to a higher level. On a national level, regional mental health providers should give priority to reaching consensus about uniform implementation of ROM. We believe that consensus is important to be able to respond to demands of policy makers. A long-term vision is necessary to achieve this goal.

References

- Anderson, K.M., Odell, P.M., Wilson, P.W., Kannel, W.B., 1991. Cardiovascular disease risk profiles. *American Heart Journal* 121, 293-298.
- Beck, A.T., Steer, R.A., Carbin, M.G., 1988. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review* 8, 77-100.
- de Beurs, E., den Hollander-Gijsman, M.E., van Rood, Y.R., van der Wee, N.J., Giltay, E.J., van Noorden, M.S., van der Lem, R., van Fenema, E.M., 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 and Psychotherapy* 18, 1-12.
- Carlier, I.V., Meuldijk, D., van Vliet, I.M., van Fenema, E., van der Wee, N.J., Zitman, F.G., 2010. Routine outcome monitoring and feedback on physical or mental health status: evidence and theory. *Journal of Evaluation in Clinical Practice* (in press).
- Carlier, I.V., Schulte-van Maaren, Y.W., Wardenaar, K.J., Giltay, E.J., van Noorden, M.S., Vergeer, P., Zitman, F.G. Development and validation of the 48-item Symptom Questionnaire in patients with depressive, anxiety and somatoform disorders. (in preparation).
- Cavanagh, J.T., Carson, A.J., Sharpe, M., Lawrie, S.M., 2003. Psychological autopsy studies of suicide: a systematic review. *Psychological Medicine* 33, 395-405.
- Conroy, R.M., Pyorala, K., Fitzgerald, A.P., Sans, S., Menotti, A., De Backer, G., De Baquer, D., Ducimetiere, P., Jousilahti, P., Keil, U., Njolstad, I., Oganov, R.G., Thomsen, T., Tunstall-Pedoe, H., Tverdal, A., Wedel, H., Whincup, P., Wilhelmsen, L., Graham, I.M., 2003. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *European Heart Journal* 24, 987-1003.
- Coryell, W., Young, E.A., 2005. Clinical predictors of suicide in primary major depressive disorder. *Journal of Clinical Psychiatry* 66, 412-417.
- Dawber, T.R., Meadors, G.F., Moore, F.E., Jr., 1951. Epidemiological approaches to heart disease: the Framingham Study. *American Journal of Public Health and the Nation's Health* 41, 279-281.
- Derogatis, L.R., Melisaratos, N., 1983. The Brief Symptom Inventory: an introductory report. *Psychological Medicine* 13, 595-605.
- Ernst, C., Lalovic, A., Lesage, A., Seguin, M., Tousignant, M., Turecki, G., 2004. Suicide and no axis I psychopathology. *BMC Psychiatry* 4, 7.
- Fava, G.A., Rafanelli, C., Tomba, E., 2011 The clinical process in psychiatry: A clinimetric approach. *Journal of Clinical Psychiatry* (in press).
- Frank, E., Cassano, G.B., Rucci, P., Thompson, W.K., Kraemer, H.C., Fagiolini, A., Maggi, L., Kupfer, D.J., Shear, M.K., Houck, P.R., Calugi, S., Grochocinski, V.J., Scocco, P., Battenfeld, J., Forgiione, R.N., 2011. Predictors and moderators of time to remission of major depression with interpersonal psychotherapy and SSRI pharmacotherapy. *Psychological Medicine* 41, 151-162.
- Gatta, G., Capocaccia, R., Stiller, C., Kaatsch, P., Berrino, F., Terenziani, M., 2005. Childhood cancer survival trends in Europe: a EUROCARE Working Group study. *Journal of Clinical Oncology* 23, 3742-3751.
- Hawton, K., van Heeringen, K., 2009. Suicide. *Lancet* 373, 1372-1381.
- Iglehart, J.K., 2009. Prioritizing comparative-effectiveness research--IOM recommendations. *New England Journal of Medicine* 361, 325-328.
- Judd, L.L., 1997. The clinical course of unipolar major depressive disorders. *Archives of General Psychiatry* 54, 989-991.

- Kadouri, A., Corruble, E., Falissard, B., 2007. The improved Clinical Global Impression Scale (iCGI): development and validation in depression. *BMC Psychiatry* 7, 7.
- Kannel, W.B., Feinleib, M., McNamara, P.M., Garrison, R.J., Castelli, W.P., 1979. An investigation of coronary heart disease in families. The Framingham offspring study. *American Journal of Epidemiology* 110, 281-290.
- Kessler, R.C., Berglund, P., Borges, G., Nock, M., Wang, P.S., 2005. Trends in suicide ideation, plans, gestures, and attempts in the United States, 1990-1992 to 2001-2003. *Journal of the American Medical Association* 293, 2487-2495.
- Lee, S.J., Earle, C.C., Weeks, J.C., 2000. Outcomes research in oncology: history, conceptual framework, and trends in the literature. *Journal of the National Cancer Institute* 92, 195-204.
- van der Lem, R., van der Wee, N.J., van Veen, T., Zitman, F.G., 2010. The generalisability of antidepressant efficacy trials to routine psychiatric out-patient practice. *Psychological Medicine* 41, 1353-1363.
- Mane, K.K., Bizon, C., Schmitt, C., Owen, P., Burchett, B., Pietrobon, R., Gersing, K., 2011. VisualDecisionLinc: A visual analytics approach for comparative effectiveness-based clinical decision support in psychiatry. *Journal of Biomedical Informatics* (in press).
- Möller, H.J., 2009. Standardised rating scales in psychiatry: methodological basis, their possibilities and limitations and descriptions of important rating scales. *World Journal of Biological Psychiatry* 10, 6-26.
- Montgomery, S.A., 1979. A new depression scale designed to be sensitive to change. *The British Journal of Psychiatry* 134, 382-389.
- Mulder, C.L., Van der Gaag, M., Bruggeman, R., Cahn, W., Delespaul, P.A., Dries, P., Faber, G., de Haan, L., van der Heijden, F.M., Kempen, R.W., Mogendorff, E.S., Slooff, C.J., Sytema, S., Wiersma, D., Wunderink, L., van Os, J., 2010. [Routine Outcome Monitoring for patients with severe mental illness: a consensus document]. *Tijdschrift voor Psychiatrie* 52, 169-179.
- Parker, G., Roy, K., Hadzi-Pavlovic, D., Mitchell, P., Wilhelm, K., 2003. Distinguishing early and late onset non-melancholic unipolar depression. *Journal of Affective Disorders* 74, 131-138.
- Penninx, B.W., Nolen, W.A., Lamers, F., Zitman, F.G., Smit, J.H., Spinhoven, P., Cuijpers, P., de Jong, P.J., van Marwijk, H.W., van der Meer, K., Verhaak, P., Laurant, M.G., de Graaf, R., Hoogendijk, W.J., van der Wee, N.J., Ormel, J., van Dyck, R., Beekman, A.T., 2011. Two-year course of depressive and anxiety disorders: Results from the Netherlands Study of Depression and Anxiety (NESDA). *Journal of Affective Disorders* (in press).
- Pui, C.H., Evans, W.E., 2006. Treatment of acute lymphoblastic leukemia. *New England Journal of Medicine* 354, 166-178.
- Pui, C.H., Sandlund, J.T., Pei, D., Campana, D., Rivera, G.K., Ribeiro, R.C., Rubnitz, J.E., Razzouk, B.I., Howard, S.C., Hudson, M.M., Cheng, C., Kun, L.E., Raimondi, S.C., Behm, F.G., Downing, J.R., Relling, M.V., Evans, W.E., 2004. Improved outcome for children with acute lymphoblastic leukemia: results of Total Therapy Study XIII at St Jude Children's Research Hospital. *Blood* 104, 2690-2696.
- Purcell, S.M., Wray, N.R., Stone, J.L., Visscher, P.M., O'Donovan, M.C., Sullivan, P.F., Sklar, P., 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 460, 748-752.
- Schulte-van Maaren, Y.W., Carlier, I.V., Giltay, E.J., van Noorden, M.S., de Waal, M.W., van der Wee, N.J., Zitman, F.G., 2011 Reference values for mental health assessment instruments: objectives and methods of the Leiden Routine Outcome Monitoring Study (in press).

- Souery, D., Oswald, P., Massat, I., Bailer, U., Bollen, J., Demyttenaere, K., Kasper, S., Lecrubier, Y., Montgomery, S., Serretti, A., Zohar, J., Mendlewicz, J., 2007. Clinical factors associated with treatment resistance in major depressive disorder: results from a European multicenter study. *Journal Clinical Psychiatry* 68, 1062-1070.
- Sox, H.C., Greenfield, S., 2009. Comparative effectiveness research: a report from the Institute of Medicine. *Annals of Internal Medicine* 151, 203-205.
- Uher, R., Perroud, N., Ng, M.Y., Hauser, J., Henigsberg, N., Maier, W., Mors, O., Placentino, A., Rietschel, M., Souery, D., Zagar, T., Czerski, P.M., Jerman, B., Larsen, E.R., Schulze, T.G., Zobel, A., Cohen-Woods, S., Pirlo, K., Butler, A.W., Muglia, P., Barnes, M.R., Lathrop, M., Farmer, A., Breen, G., Aitchison, K.J., Craig, I., Lewis, C.M., McGuffin, P., 2010. Genome-wide pharmacogenetics of antidepressant response in the GENDEP project. *The American Journal of Psychiatry* 167, 555-564.
- Veen, G., van Vliet, I.M., De Rijk, R.H., Giltay, E.J., van Pelt, J., Zitman, F.G., 2011. Basal cortisol levels in relation to dimensions and DSM-IV categories of depression and anxiety. *Psychiatry Research* 185, 121-128.
- Vuorilehto, M.S., Melartin, T.K., Isometsa, E.T., 2009. Course and outcome of depressive disorders in primary care: a prospective 18-month study. *Psychological Medicine* 39, 1697-1707.
- Wardenaar, K.J., Vreeburg, S.A., van Veen, T., Giltay, E.J., Veen, G., Penninx, B.W., Zitman, F.G., 2011. Dimensions of depression and anxiety and the hypothalamo-pituitary-adrenal axis. *Biological Psychiatry* 69, 366-373.
- Washington, A.E., Lipstein, S.H., 2011. The Patient-Centered Outcomes Research Institute - Promoting Better Information, Decisions, and Health. *New England Journal of Medicine* (in press).
- World Health Organisation. World report on violence and health. 2002, WHO edn., Geneva.
- Zimmerman, M., Posternak, M.A., Chelminski, I., Friedman, M., 2005. Standardized clinical outcome rating scale for depression for use in clinical practice. *Depression and Anxiety* 22, 36-40.
- Zisook, S., Lesser, I., Stewart, J.W., Wisniewski, S.R., Balasubramani, G.K., Fava, M., Gilmer, W.S., Dresselhaus, T.R., Thase, M.E., Nierenberg, A.A., Trivedi, M.H., Rush, A.J., 2007. Effect of age at onset on the course of major depressive disorder. *The American Journal of Psychiatry* 164, 1539-1546.
- Zisook, S., Rush, A.J., Albala, A., Alpert, J., Balasubramani, G.K., Fava, M., Husain, M., Sackeim, H., Trivedi, M., Wisniewski, S., 2004. Factors that differentiate early vs. later onset of major depression disorder. *Psychiatry Research* 129, 127-140.

Chapter 8

Nederlandse samenvatting

List of abbreviations

List of publications

Curriculum vitae

Acknowledgements (dankwoord)

Nederlandse samenvatting

Dit proefschrift gaat over het routinematig meten van klachten bij patiënten met Stemnings-Angst-en-Somatoforme (SAS) stoornissen met behulp van gestandaardiseerde meetinstrumenten. Allereerst worden de klachten, symptomen en diagnoses systematisch vastgesteld, nog voordat een behandeling is gestart, om de beginsituatie vast te leggen. Vervolgens worden door vervolgmetingen behandeluitkomsten in kaart gebracht. De Engelse term van het routinematig meten van klachten is 'Routine Outcome Monitoring', afgekort 'ROM'. Het routinematig karakter van ROM maakt dat de metingen in grote mate representatief zijn voor de dagelijkse behandelpraktijk ('real-world'). SAS stoornissen zijn potentieel invaliderende ziekten die frequent voorkomen. Zo krijgt ongeveer één op de vijf Nederlanders in zijn of haar leven een depressieve of angststoornis. De meeste patiënten worden in de eerste lijn behandeld, door de huisarts of door een eerstelijnspsycholoog. Als de klachten ernstiger zijn of niet reageren op eenvoudige interventies, vindt vaak doorverwijzing plaats naar de tweede lijn: de Geestelijke Gezondheidszorg (GGZ). Bij onvoldoende behandelresultaat of complexe problematiek kan de patiënt naar de derde lijn verwezen worden; bijvoorbeeld een universitair medisch centrum. Deze manier van behandelen, die overigens niet specifiek is voor de psychiatrie, heet 'stepped care'. De studies in dit proefschrift zijn gebaseerd op behandelgegevens van patiënten in de tweede en derde lijn.

In de dagelijkse behandelpraktijk in de tweede en derde lijn worden patiënten behandeld vanwege een breed scala aan klachten. Vaak bestaan er naast de SAS-stoornis waarvoor de patiënt hulp zoekt nog andere stoornissen ('comorbiditeit'). Zo kan er bijvoorbeeld ook sprake zijn van middelenmisbruik, of een persoonlijkheidsstoornis. Patiënten met een depressie zijn vaak ook angstig, en patiënten met een angststoornis vaak ook somber. Daarnaast kan er sprake zijn van gebrekkige therapietrouw waardoor de behandeling minder effectief wordt. Deze en andere factoren spelen ook een rol bij het kiezen van de behandeling, en kunnen maken dat in de dagelijkse behandelpraktijk wordt afgeweken van reguliere behandelprotocollen of -richtlijnen.

Veel kennis over kenmerken van psychiatrische ziekten is afkomstig uit bevolkingsonderzoeken. Bij dit type onderzoek worden de klachten van een grote steekproef van mensen uit bijvoorbeeld een land in kaart gebracht door middel van gestandaardiseerde meetinstrumenten. Een groot deel van deze mensen heeft klachten zonder dat een behandeling gezocht wordt. Ook is de laatste decennia veel kennis over kenmerken en behandelresultaten van SAS-stoornissen afkomstig van gerandomiseerde gecontroleerde onderzoeken ('Randomised Controlled Trials', afgekort 'RCT's'). Bij RCT's is

vaak de primaire vraag of een behandeling 'x' beter werkt dan een behandeling 'y'. Als 'bijzaak' van RCT's is veel kennis opgedaan over de kenmerken van de ziekte zelf. Bij dit type onderzoek worden vaak strenge selectiecriteria voor inclusie gehanteerd. Behandelingen volgens strakke protocollen worden uitgevoerd. Zo worden patiënten met comorbiditeit of middelenmisbruik vaak uitgesloten van RCT's. Het is aangetoond dat het grootste deel van de patiënten in de algemene behandelpraktijk om deze redenen niet zouden kunnen participeren in een RCT ('selectiebias' van RCT's). Bevindingen van deze RCT's, maar ook die van bevolkingsonderzoeken zijn dus waarschijnlijk slechts beperkt representatief voor de patiënten uit de dagelijkse behandelpraktijk. Vanwege het routinematig karakter van ROM-metingen bij patiënten die behandeling zoeken, zouden deze metingen gegevens op kunnen leveren over kenmerken van ziektes en behandelresultaten die in grote mate representatief zijn voor de klinische praktijk.

In Leiden en omstreken bestaat sinds 2001 een ROM-infrastructuur waarbij patiënten die verwezen worden voor behandeling van een SAS-stoornis naar een polikliniek van GGZ Rivierduinen (RD) of de afdeling psychiatrie van het Leids Universitair Medisch Centrum (LUMC) tijdens de intake een ROM meting krijgen. Alleen patiënten die te ziek zijn om een ROM te ondergaan, of die de Nederlandse taal onvoldoende machtig zijn, worden uitgesloten van ROM. Als een behandeling wordt ingesteld, vindt elke 3-4 maanden een ROM-vervolgmeting plaats. ROM-metingen worden uitgevoerd door speciaal getrainde testverpleegkundigen en zijn primair bedoeld voor de arts en patiënt, voor diagnostiek en om voortgang van de behandeling te kunnen evalueren. ROM gegevens zijn echter ook geanonimiseerd beschikbaar voor wetenschappelijk onderzoek. Zo kunnen van grote groepen patiënten uit de dagelijkse behandelpraktijk ('real-world patients') gegevens over klachten, symptomen, diagnoses en behandel effecten gebruikt worden voor het in kaart brengen van karakteristieken van ziektebeelden of het identificeren van voorspellers van behandelresultaten in de dagelijkse behandelpraktijk ('real-world outcomes'). Daarnaast kunnen ROM gegevens gebruikt worden voor 'benchmarking': het vergelijken van behandelresultaten van instituten, afdelingen of individuele behandelaren voor managementdoeleinden.

Doelen van de studies opgenomen in dit proefschrift waren ten eerste het beschrijven van correlaten van bepaalde ziektekenmerken bij grote groepen van patiënten met SAS stoornissen in de dagelijkse behandelpraktijk, en ten tweede het identificeren van voorspellers van behandeluitkomsten in de dagelijkse behandelpraktijk. Een secundair doel was het nagaan van de bruikbaarheid van ROM data voor wetenschappelijk onderzoek.

Belangrijkste bevindingen

In **hoofdstuk 1** geven wij allereerst een overzicht van de achtergronden van ROM, waarbij het belang van objectieve en gestandaardiseerde metingen voor psychiatrische diagnostiek duidelijk naar voren komt. Psychiatrische ziekten worden geclassificeerd met behulp van de DSM-IV: de vierde editie van de 'Diagnostic and Statistical Manual of Mental Disorders'. Dit is een internationaal classificatiesysteem waarin de criteria van alle psychiatrische ziekten beschreven zijn. Omdat ondanks de sterk toegenomen kennis van het brein de onderliggende biologische mechanismen van veel psychiatrische ziekten tot op heden niet bekend zijn moeten klachten en symptomen van de patiënt gebruikt worden voor diagnostische classificatie. Om dit zorgvuldig te kunnen doen, is het van belang om deze klachten en symptomen zo objectief mogelijk vast te stellen. Dat is vaak niet eenvoudig omdat veel psychische symptomen een grote subjectieve component hebben: bijvoorbeeld verstoorde stemming, angst of wanen. Het gebruik van gevalideerde, gestandaardiseerde meetinstrumenten bij de psychiatrische diagnostiek verhoogt in belangrijke mate de betrouwbaarheid van deze diagnostiek. De ontwikkeling enkele decennia geleden van betrouwbare diagnostische meetinstrumenten die gebruik maken van DSM-criteria heeft de mogelijkheden voor (epidemiologisch) wetenschappelijk onderzoek enorm vergroot. Dit heeft geresulteerd in bevolkingsonderzoeken waarin de prevalentie (het vóórkomen) van psychiatrische ziekten kon worden onderzocht. Ook zijn vele meetinstrumenten ontwikkeld, die de ernst van bepaalde ziekten (bijvoorbeeld depressie) meten op symptoomniveau, of door het sociaal functioneren in kaart te brengen. Voorbeelden van deze lijsten zijn de Montgomery Asberg Depression Rating Scale (MADRS; een observatielijst die de ernst van depressie meet) en de Short-Form-36 (SF-36; een zelfrapportagelijst die de algemene gezondheidstoestand en sociaal functioneren meet). Deze en andere lijsten kunnen bij ROM worden gebruikt om het verloop van klachten over de tijd te meten.

Vervolgens geven wij in dit hoofdstuk een overzicht van ROM in het algemeen, en ROM in RD en LUMC, en een beschrijving van de ROM-patiëntenpopulatie. Daarnaast zetten wij de doelen van het proefschrift uiteen.

In de eerste studie beschreven in dit proefschrift (**hoofdstuk 2**) bestudeerden wij geslachtsverschillen bij patiënten met een 'Major' depressieve stoornis (MDD). Hoewel geslachtsverschillen in het beloop en symptoomprofiel van MDD al vaak zijn bestudeerd, blijkt dat bijna alle studies zijn verricht in de algemene bevolking, dus bij mensen die niet noodzakelijkerwijs in behandeling zijn, of in RCT-populaties. Gegevens van 'real-world' patiënten waren nauwelijks beschikbaar. De bevindingen uit eerdere

studies waren niet altijd consistent. Dit is mogelijk te verklaren door de selectieve populaties die waren gebruikt voor deze studies. Toch wordt algemeen aangenomen dat er belangrijke verschillen bestaan tussen mannen en vrouwen met MDD, in termen van voorkomen, beloop en symptoompatronen. Doel van deze studie was het onderzoeken van geslachtsverschillen in klinische karakteristieken bij 'real-world' patiënten met MDD. Wij gebruikten de eerste ROM meting ('baseline') van 1131 patiënten met MDD. De belangrijkste bevindingen van deze studie waren dat vrouwen meer en ernstiger klachten rapporteerden op 'subjectieve' zelfrapportagelijsten, en een lagere algemene gezondheidsbeleving hadden dan mannen. Op meer objectieve observatielijsten vonden wij geen verschillen in ernst van symptomen tussen mannen en vrouwen. Deze bevindingen suggereren dat de subjectieve ziektelast hoger is bij vrouwen met MDD dan bij mannen met MDD. Ook vonden wij dat het symptoompatroon tussen mannen en vrouwen verschilde, waarbij vrouwen met MDD meer verlies van energie, slaperigheid, gevoelens van waardeloosheid, verlies van eetlust, moeheid en libidoverlies rapporteerden dan mannen. Deze bevindingen kwamen redelijk overeen met eerder gerapporteerde bevindingen van RCT's of populatiestudies. Een belangrijke conclusie van deze eerste studie was dat voor een evenwichtig en betrouwbaar beeld de ROM dient te bestaan uit zowel zelfrapportage- als observatielijsten.

In **hoofdstuk 3** gebruikten wij ook eerste ROM metingen van patiënten met MDD. In deze studie onderzochten wij verschillen in klinische karakteristieken tussen patiënten met een MDD ontstaan voor het 18e jaar en patiënten met een MDD ontstaan na het 18e jaar. Eerdere studies vonden aanwijzingen dat de op vroege leeftijd ontstane MDD mogelijk een ernstigere vorm is van dezelfde ziekte. Ook deze studies waren echter weer voornamelijk in RCT-populaties en in de algemene bevolking verricht. Wij gebruikten een bestand van 1105 patiënten met een MDD en een eerste ROM meting, om te onderzoeken of deze verschillen ook aangetoond konden worden bij 'real-world' patiënten. De belangrijkste bevinding van deze studie was dat patiënten met een vroeg ontstane depressie vaker een suïcidepoging in het verleden hadden gedaan, en ook vaker suïcidale gedachten hadden op het moment van de ROM meting dan patiënten met een depressie ontstaan na het 18e jaar. In onze analyses vonden wij geen verschillen op andere gebieden van klachten, of in ernst van de symptomen, verschillen die in veel eerdere studies wel gevonden werden. De leeftijd van ontstaan kan niet altijd betrouwbaar worden vastgesteld omdat de patiënt dit misschien niet meer kan reproduceren, toch kan het de clinicus op een relatief eenvoudige wijze helpen bij het inschatten van suïcidaliteit bij patiënten met MDD.

Na deze twee studies bij patiënten met MDD gebruikten wij in de volgende studie (**hoofdstuk 4**) de gehele groep van patiënten met SAS stoornissen voor analyses. In deze studie bestudeerden wij met behulp van de eerste ROM meting de prevalentie en factoren die samenhangen met opzettelijke zelfbeschadiging en suïcidale ideaties (DSHI) in een groep van 2844 patiënten die verwezen waren naar RD of LUMC voor behandeling van een SAS-stoornis. Uit eerder onderzoek bleek dat DSHI een belangrijke voorspeller is voor een geslaagde suïcide. In deze studie vonden wij dat opzettelijke zelfbeschadiging bij 55% van de patiënten voorkwam. Getrouwd zijn, hogere opleiding en hogere angstscores waren onafhankelijk geassocieerd met een lagere kans op DSHI. Tegelijkertijd voorkomende SAS-diagnoses, het hebben van meer depressieve symptomen en emotieregulatiestoornissen hingen duidelijk samen met een hogere kans op DSHI. Deze resultaten in een grote groep 'real-world' patiënten met SAS-stoornissen laten zien dat DSHI vaak voorkomt bij patiënten met SAS-stoornissen, en dat er duidelijke en gemakkelijk vast te stellen factoren zijn die ermee samenhangen. De bevindingen uit deze studie kunnen klinici helpen bij het identificeren van patiënten met een verhoogd risico op DSHI en suïcide, en hier in hun behandeling speciaal aandacht aan besteden.

In het volgende gedeelte van dit proefschrift gebruikten wij naast eerste ROM-metingen ook de vervolgmetingen. Hierdoor was het mogelijk om voorspellers van behandeluitkomsten te onderzoeken. In **hoofdstuk 5** gebruikten wij van 1489 patiënten met een MDD de eerste ROM-meting en vervolgmetingen tot 24 maanden na de eerste meting. Het doel van deze studie was te onderzoeken of de individuele depressieve symptomen gemeten op een zelfrapportagelijst (Beck Depression Inventory-Revised Version; BDI-II) de behandeluitkomst konden voorspellen. De BDI-II bestaat uit 21 vragen die gaan over afzonderlijke depressieve symptomen, waarbij de patiënt moet aangeven in hoeverre hij de afgelopen twee weken last heeft gehad van het betreffende symptoom. In eerdere studies naar voorspellers van uitkomsten van depressie werd al gevonden dat klinische factoren zoals bijvoorbeeld een vroege leeftijd van ontstaan, ernstiger klachten en bepaalde persoonlijkheidstrekken geassocieerd waren met een slechtere behandeluitkomst. Het is belangrijk om al in een vroeg stadium patiënten met een hoog risico op een slechte behandeluitkomst te kunnen identificeren omdat deze patiënten mogelijk baat zouden kunnen hebben bij andere behandelingen dan de algemene voorgeschreven behandelingen. Zo valt bijvoorbeeld te denken aan extra psychotherapeutische interventies bij deze groep. Het systematisch onderzoeken of bepaalde depressieve symptomen bij aanvang van de behandeling ook een voorspellende waarde zouden hebben, was nog niet eerder gedaan. Wij vonden dat 2 van de 21 symptomen sterke onafhankelijke voorspellers waren van een slechte behandeluitkomst:

patiënten die hoog scoorden op de symptomen ‘pessimisme’ en ‘verlies van energie’ hadden meer kans op een slechte behandeluitkomst. Deze uitkomsten waren gecorrigeerd voor eerder gevonden correlaten van slechte behandeluitkomst, en kwamen overeen met de bevindingen van een recente studie waarbij de voorspellende waarde van *clusters* van symptomen (in plaats van afzonderlijke symptomen, zoals in onze studie) bij depressie was onderzocht. Wij achten deze bevindingen voor de clinicus erg relevant, omdat deze voorspellende symptomen gemakkelijk en goedkoop zijn vast te stellen (bijvoorbeeld door het afnemen van een zelfrapportagelijst) voor aanvang van de behandeling.

De laatste studie beschreven in dit proefschrift (**hoofdstuk 6**) behandelt eveneens een onderzoek naar voorspellers van behandeluitkomst in de dagelijkse behandelpraktijk, maar nu in de gehele groep van patiënten met SAS-stoornissen. SAS-stoornissen vertonen een grote mate van overlap in diagnostische criteria, en vaak zijn er meerdere comorbide SAS-stoornissen aanwezig zonder dat duidelijk is welke van deze stoornissen op de voorgrond staat. Toch worden in de meeste RCT's patiënten met comorbiditeit uitgesloten van deelname aan deze studies, hetgeen de generaliseerbaarheid van de resultaten beperkt. Wij gebruikten eerste ROM metingen en vervolgmetingen tot 24 maanden na de eerste meting van twee cohorten met in totaal 2284 patiënten met één of meer SAS-stoornissen, en onderzochten welke karakteristieken een slechte behandeluitkomst voorspelden. Wij vonden dat een hogere leeftijd, het hebben van meerdere SAS-stoornissen of het hebben van een somatoforme stoornis, en bepaalde persoonlijkheidstrekken (problemen met intimiteit, affectabiliteit, en de neiging tot zelfbeschadiging), onafhankelijke voorspellers waren voor een slechte behandeluitkomst. In toekomstige studies zou onderzocht moeten worden of deze patiënten baat kunnen hebben bij extra of andere interventies.

Algemene discussie

De studies in dit proefschrift hebben inzicht gegeven in een aantal klinische aspecten van ‘real-world’ patiënten met SAS-stoornissen. Daarnaast zijn een aantal voorspellers voor een slechte behandeluitkomst vastgesteld. Deze voorspellers zijn eenvoudig te meten en direct klinisch toepasbaar. De studies in dit proefschrift illustreren eveneens de mogelijkheden van wetenschappelijk onderzoek met ROM data. Ook maken deze studies ons bewust van de mogelijke beperkingen van ROM.

Toekomstige mogelijkheden van wetenschappelijk onderzoek met ROM data zijn bijvoorbeeld het nauwkeurig uitwerken van risicofactoren op slechte

behandeluitkomsten bij bepaalde groepen van patiënten. Dit kan resulteren in tabellen met risicoschattingen, zoals bijvoorbeeld ook al veelvuldig gebruikt worden bij hart- en vaatziekten (zie Hoofdstuk 7; figuur 7.1). In vervolgonderzoek zou als volgende stap gekeken kunnen worden naar het effect van extra of andere interventies bij subgroepen van patiënten met een verhoogd risico op een slechte behandelrespons. Dit is een voorbeeld van 'Comparative Effectiveness Research' (CER), en dit soort studies zou kunnen leiden tot meer geïndividualiseerde behandelingen. Om dit te realiseren is nationale en internationale samenwerking nodig. Dat dit complex maar niet onmogelijk is, blijkt bijvoorbeeld uit de oncologie: door internationale samenwerking waarbij decennia lang structureel en systematisch is gewerkt aan staging, toepassen van behandelprotocollen en monitoring van behandeluitkomsten is de vijfjaarsoverleving van kinderen met acute lymfatische leukemie spectaculair toegenomen van rond de 50% in 1980 tot meer dan 85% in 2012. Om vergelijkbare resultaten te bereiken in de psychiatrie zou op grote schaal consensus moeten worden verkregen met betrekking tot relevante ROM-metingen. Welke pakket aan meetinstrumenten optimaal is bij SAS-patiënten zou moeten worden onderzocht. Dat niet kan worden volstaan met enkele zelfrapportagelijsten is waarschijnlijk, omdat voor een gebalanceerde ROM-meting zowel zelfrapportage- als observatielijsten nodig zijn (Hoofdstuk 2). Gedetailleerde informatie over symptomen, comorbiditeit, psychosociaal functioneren, kwaliteit van leven verkregen met ROM zou een belangrijke basis kunnen vormen van staging van SAS-patiënten en het evalueren van behandeluitkomsten.

Mogelijke beperkingen van ROM onderzoek zijn aan het licht gekomen door het verrichten van de studies in dit proefschrift. Allereerst zijn er de beperkingen van de huidige ROM-infrastructuur van RD en LUMC. De belangrijkste beperking is het ontbreken van een koppeling tussen het elektronische patiëntendossier (EPD) en de ROM. Hierdoor zijn geen gedetailleerde behandelgegevens beschikbaar, maar ook geen gegevens over bijvoorbeeld psychiatrische voorgeschiedenis en familieanamnese. Hierdoor hebben wij in de prospectieve studies niet kunnen corrigeren voor het type behandeling of de mate waarin richtlijnen gevolgd zijn.

Een belangrijke mogelijke beperking van ROM in het algemeen is dat het meten van klachten en symptomen door middel van meetinstrumenten altijd onnauwkeurigheden met zich mee brengt. Om een zo betrouwbaar mogelijk beeld van de klachten van de patiënt te krijgen, is het noodzakelijk om een brede testbatterij te gebruiken, die bestaat uit zowel zelfrapportagelijsten als observatieschalen, en die verschillende domeinen bestrijkt. In tijden van bezuiniging en kostenbeheersing van de zorg is het erg belangrijk om behandelresultaten te meten, om deze te kunnen vergelijken ('benchmarking'). Inmiddels

is ROM verplicht gesteld en het zal niet lang meer duren voordat GGZ-instellingen worden afgerekend op behandelresultaten gemeten met ROM. Tegelijkertijd bestaat de neiging om dure observatielijsten weg te laten en een zo beknopt mogelijke ROM te realiseren. Dat hier een potentiële verarming op de loer ligt moge duidelijk zijn. Bovendien neemt het risico op misinterpretatie van ROM gegevens en daaruit voorkomende ongefundeerde oordelen toe naarmate de ROM metingen beknopter worden. Bijvoorbeeld: in hoofdstuk zes vonden wij dat bepaalde dysfunctionele persoonlijkheidstrekken voorspellend waren voor een slechtere behandeluitkomst. Het meten van deze persoonlijkheidstrekken kost behoorlijk wat tijd. Bij het weglaten van het betreffende instrument zou deze informatie verloren gaan, en zouden afdelingen met meer ‘moeilijke’ patiënten onterecht kunnen worden afgerekend op tegenvallende behandelresultaten.

Perspectieven voor toekomstig onderzoek

De studies in dit proefschrift kunnen worden beschouwd als exploratieve studies om het potentieel van epidemiologisch onderzoek met ROM-data te illustreren. Een aantal aanpassingen, zoals de koppeling van het EPD aan de ROM, zal de mogelijkheden van onderzoek met ROM-data in RD en LUMC aanzienlijk vergroten. Naast het dienen als basis voor epidemiologisch onderzoek, wordt de RD en LUMC ROM database inmiddels als basis gebruikt voor onderzoek in andere domeinen. Voorbeelden van deze domeinen zijn biologisch en psychometrisch onderzoek (zie tabel 7.1) en het gebruik van de ROM infrastructuur voor het verrichten van metingen in RCTs.

De grootste uitdaging voor de toekomst is het vaststellen van een optimale testbatterij en ROM-infrastructuur voor verschillende behandelsettings en patiëntencategorieën, waarbij het bereiken van (inter)nationale consensus (ten aanzien van de gebruikte zelfrapportage- en observatielijsten en de bijbehorende norm-waardes) de mogelijkheden van vooruitgang enorm zal vergroten.

List of abbreviations

ADHD	Attention Deficit Hyperactivity Disorder
AN(C)OVA	Analysis of (co)variance
APA	American Psychiatric Association
BAS	Brief Anxiety Scale
BDI-II	Beck Depression Inventory, Revised Edition
BSI	Brief Symptom Inventory
CER	Comparative Effectiveness Research
CGI-S	Clinical Global Impression-Severity
CI	Confidence Interval
CIDI	Composite International Diagnostic Interview
COROM	ROM Research Center Leiden
CPRS	Comprehensive Psychopathological Rating Scale
CVD	Cardiovascular Disease
DAPP-SF	Dimensional Assessment of Personality Pathology-Short Form
DIS	Diagnostic Interview Schedule
DSHI	Deliberate Self Harm and Suicidal Ideation
DSM(-IV TR)	Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition, Text Revision)
ECA	Epidemiological Catchment Area
GAF	Global Assessment of Functioning
HDL	High-Density Lipoprotein
HDRS	Hamilton Depression Rating Scale
HPA-axis	Hypothalamus Pituitary Adrenal-axis
HR	Hazard Ratio
IQR	Interquartile Range
LUMC	Leiden University Medical Center
MADRS	Montgomery-Åsberg Depression Rating Scale
MAS	Mood, Anxiety and Somatoform
MASHBANK	Mood, Anxiety, Somatoform disorders and HPA-axis Biobank
MASQ	Mood and Anxiety Symptom Questionnaire
MDD	Major Depressive Disorder
MDE	Major Depressive Episode
MEC	Medical Ethical Committee

MINI-Plus	Mini International Neuropsychiatric Interview-Plus
NCS(-R)	National Comorbidity Survey(-Replication)
NEMESIS	Netherlands Mental Health Survey and Incidence Study
NIMH	National Institute of Mental Health
OR	Odds Ratio
PAREL	Psychiatrische Academische Registratie Leiden
PCORI	Patient-Centered Outcomes Research Institute
PD	Panic Disorder
PRF	Patient Record Form
QoL	Quality of Life
RCT	Randomised Controlled Trial
RD	Rivierduinen
RMHP	Regional Mental Health Provider
ROM	Routine Outcome Monitoring
SBP	Systolic Blood Pressure
SCL-90	Symptom Checklist-90
SCORE	Systematic Cardiovascular Risk Evaluation
SD	Standard Deviation
SF-36	Short Form-36
SMI	Severe Mental Illness
SPSS	Statistical Package for the Social Sciences
SQ-48	Symptom Questionnaire-48
STAR*D	Sequenced Treatment Alternatives to Relieve Depression
US	United States
WHO	World Health Organisation

List of publications

van Noorden, M.S.*, van Fenema, E.M.*, van der Wee, N.J.A., van Rood, Y.R., Carlier, I.V.E., Zitman, F.G., Giltay, E.J. Predicting outcomes of mood, anxiety and somatoform disorders: the Leiden Routine Outcome Monitoring Study. *Journal of Affective Disorders* (in press).

*** shared first authorship**

van Noorden, M.S., van der Wee, N.J.A., Zitman, F.G., Giltay E.J. Routine outcome monitoring in psychiatric clinical practice: backgrounds and overview. *The International Journal of Person Centered Medicine* (in press).

van Amsterdam, J.G.C., Brunt, T.M., McMaster, M.T.B., Niesink, R., **van Noorden, M.S.**, van den Brink, W. Cognitieve schade door intensief gebruik en overdoses van GHB. *Tijdschrift voor Psychiatrie* (in press).

van Noorden, M.S., van Fenema, E.M., van der Wee, N.J., Zitman, F.G., Giltay, E.J., 2012. Predicting outcome of depression using the depressive symptom profile: the Leiden Routine Outcome Monitoring Study. *Depression and Anxiety*. Epub ahead of print.

Schulte-van Maaren, Y.W., Carlier, I.V.E., Giltay, E.J., **van Noorden, M.S.**, de Waal, M.W., van der Wee, N.J.A., Zitman, F.G., 2012. Reference values for mental health assessment instruments: objectives and methods of the Leiden Routine Outcome Monitoring Study. *Journal of Evaluation in Clinical Practice*. Epub ahead of print.

van der Meer, J., van Rood, Y.R., van der Wee, N.J., den Hollander-Gijsman, M., **van Noorden, M.S.**, Giltay, E.J., Zitman, F.G., 2011. Prevalence, demographic and clinical characteristics of body dysmorphic disorder among psychiatric outpatients with mood, anxiety or somatoform disorders. *Nordic Journal of Psychiatry*. Epub ahead of print.

de Rover, M., Brown, S.B., Boot, N., Hajcak, G., **van Noorden, M.S.**, van der Wee, N.J., Nieuwenhuis, S., 2012. Beta receptor-mediated modulation of the late positive potential in humans. *Psychopharmacology (Berl)* 219, 971-979.

de Klerk, S.*, **van Noorden, M.S.***, van Giezen, A.E., Spinhoven, P., den Hollander-Gijsman, M.E., Giltay, E.J., Speckens, A.E., Zitman, F.G., 2011. Prevalence and correlates of lifetime deliberate self-harm and suicidal ideation in naturalistic outpatients: the Leiden Routine Outcome Monitoring study. *Journal of Affective Disorders* 133, 257-264. * **Shared first authorship.**

van der Loos, M.L., Mulder, P., Hartong, E.G., Blom, M.B., Vergouwen, A.C., **van Noorden, M.S.**, Timmermans, M.A., Vieta, E., Nolen, W.A.; LamLit Study Group., 2011. Long-term outcome of bipolar depressed patients receiving lamotrigine as add-on to lithium with the possibility of the addition of paroxetine in nonresponders: a randomized, placebo-controlled trial with a novel design. *Bipolar Disorders* 13, 111-117.

van Noorden, M.S., Minkenberg, S.E., Giltay, E.J., den Hollander-Gijsman, M.E., van Rood, Y.R., van der Wee, N.J., Zitman, F.G., 2011. Pre-adult versus adult onset major depressive disorder in a naturalistic patient sample: the Leiden Routine Outcome Monitoring Study. *Psychological Medicine* 41, 1407-1417.

van Noorden, M.S., Kamal, R., de Jong, C.A., Vergouwen, A.C., Zitman, F.G., 2010. [Gamma-hydroxybutyric acid (GHB) dependence and the GHB withdrawal syndrome: diagnosis and treatment]. *Nederlands Tijdschrift voor Geneeskunde* 154, A1286.

de Rover, M., **van Noorden, M.S.**, Nieuwenhuis, S., van der Wee, N.J., 2010. Cardiorespiratory arrest in a healthy volunteer after a single oral dose of 80 mg of the beta-blocker propranolol. *Neurobiology of Learning and Memory* 94, 576-577.

de Beurs, E., den Hollander-Gijsman, M.E., van Rood, Y.R., van der Wee, N.J., Giltay, E.J., **van Noorden, M.S.**, van der Lem, R., van Fenema, E.M., 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 and Psychotherapy* 18, 1-12.

van der Loos, M.L., Mulder, P., Hartong, E.G., Blom, M.B., Vergouwen, A.C., **van Noorden, M.S.**, Timmermans, M.A., Vieta, E., Nolen, W.A.; LamLit Study Group., 2010. Efficacy and safety of two treatment algorithms in bipolar depression consisting of a combination of lithium, lamotrigine or placebo and paroxetine. *Acta Psychiatrica Scandinavica* 122, 246-254.

van Noorden, M.S., Giltay, E.J., den Hollander-Gijsman, M.E., van der Wee, N.J., 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.

van der Wee, N.J., **van Noorden, M.S.**, Veltman, D.J., 2009. [Genetic neuroimaging]. *Tijdschrift voor Psychiatrie* 51, 665-673.

van Noorden, M.S., van Dongen, L.C., Zitman, F.G., Vergouwen, T.A., 2009. Gamma-hydroxybutyrate withdrawal syndrome: dangerous but not well-known. *General Hospital Psychiatry* 31, 394-6.

Klaassens, E.R., **van Noorden, M.S.**, Giltay, E.J., van Pelt, J., van Veen, T., Zitman, F.G., 2009. Effects of childhood trauma on HPA-axis reactivity in women free of lifetime psychopathology. *Progress in Neuropsychopharmacology and Biological Psychiatry* 33, 889-894.

van Noorden, M.S., Vergouwen, A.C., Koerselman, G.F., 2002. [Delirium during withdrawal of venlafaxine]. *Nederlands Tijdschrift voor Geneeskunde* 146, 1236-1237.

Curriculum Vitae

Martijn van Noorden is geboren op 18 januari 1975 te Amsterdam. In 1993 behaalde hij zijn gymnasiumdiploma aan het Sint Nicolaaslyceum in Amsterdam. Hierna studeerde hij geneeskunde aan de Vrije Universiteit te Amsterdam. Tijdens zijn studie verrichtte hij bestuurswerk voor de Medische Faculteitsvereniging aan de Vrije Universiteit, was hij mentor voor eerstejaarsstudenten en verrichtte hij diverse studentassistentenschappen. Hij volgde zijn wetenschappelijke stage bij de vakgroep anesthesie van het Parirenyatwa Hospital in Harare, Zimbabwe. Zijn co-schap psychiatrie volgde hij bij de crisisdienst in Amsterdam. Na het behalen van zijn artsenbul in 2000 werkte hij een jaar als assistent (AGNIO) interne geneeskunde in het Sint Lucas Andreas Ziekenhuis te Amsterdam. Vanaf 2001 volgde hij de opleiding tot psychiater in hetzelfde ziekenhuis. De stage sociale psychiatrie volgde hij bij Mentrum in Amsterdam. Zijn keuzestage volgde hij in het Leids Universitair Medisch Centrum, waar hij vanaf oktober 2006 als psychiater werkzaam is. Hij heeft ruim vijf jaar als supervisor op de polikliniek stemmingsstoornissen gewerkt. Sinds 2012 werkt hij als psychiater op de afdeling psychiatrie. Hij was betrokken bij de oprichting en uitvoering van de MASHBANK, een biobank waarin biologisch materiaal van 2000 patiënten met stemmings-, angst- en somatoforme stoornissen en gezonde controles is opgeslagen. In 2008 startte hij het promotieonderzoek dat resulteerde in het huidige proefschrift. Verder is hij betrokken bij onderwijs aan geneeskundestudenten en de opleiding van AIOS. Hij woont samen en heeft drie dochters.

Martijn van Noorden was born in 1975 in Amsterdam, the Netherlands. After completing secondary school education, he started his medical education at the Vrije Universiteit in Amsterdam. In 2000 he graduated as a medical doctor. After obtaining his medical degree he worked as an internal medicine resident in the Sint Lucas Andreas hospital in Amsterdam for one year. He then started his training as a psychiatry resident in the same hospital. He completed his psychiatric training in October 2006 and started working as a psychiatrist in the Leiden University Medical Center. He has worked as supervisor at the mood disorders outpatient department until 2012. From January 2012 he works at the inpatient department. He is involved in medical education and training of psychiatry residents. He started his PhD-project in 2008.

Acknowledgements (dankwoord)

Het schrijven van dit proefschrift was een mooi project. Nu het voltooid is past het om de mensen te bedanken die mij de afgelopen vier jaar, op wat voor manier dan ook, hebben bijgestaan.

Allereerst bedank ik mijn promotor, Frans Zitman, en beide co-promotores Erik Giltay en Nic van der Wee.

Frans, aanvankelijk was het plan dat ik zou promoveren op de MASHBANK, de biobank die je wilde oprichten om biologische gegevens van grote groepen patiënten te kunnen koppelen aan ROM-gegevens. In de eerste maanden dat ik in Leiden werkte, kwamen we overeen dat ik deze biobank zou gaan optuigen. Na twee jaar hard werken was de MASHBANK geïmplementeerd in het LUMC en diverse locaties binnen Rivierduinen, maar de inclusie verliep niet zo snel als gehoopt. Omdat ik graag wilde starten met onderzoek, bespraken we dat ik zou beginnen met ROM-analyses totdat de MASHBANK-data beschikbaar zouden komen. Uiteindelijk is het proefschrift een ROM-proefschrift geworden, en is de MASHBANK inclusie nu ook afgerond. De MASHBANK-analyses kunnen nu beginnen. Ik dank je voor je steun aan dit 'alternatieve' plan, en voor je kritische begeleiding gedurende het gehele traject.

Erik, met jou is het goed samenwerken. Als co-promotor ben je in staat om razendsnel te reageren op vragen, of het nu gaat om problemen met analyses of het becommentariëren van manuscripten, en altijd bereid om praktisch en creatief mee te denken. Ook als directe collega in de kliniek geniet ik van onze samenwerking.

Nic, jij voegde je later bij de promotiecommissie, maar was vanaf het begin al betrokken bij de studies. Ik vind het prettig om strategisch met je te overleggen en je snelle analyses, vaak op meta-niveau, van ingewikkelde problemen mee te nemen in mijn werk. Ook jij bent ondanks de steeds meer toenemende drukte een collega bij wie ik laagdrempelig terecht kan om even te 'sparren'.

Esther, bij enkele studies uit ons beider proefschriften hebben we samengewerkt. Je nuchtere aanpak vergezeld van dubbele espresso's werkte aanstekelijk. Mijn andere collega-psychiaters Roos, Jacqueline, Irene, tot voor kort Manon en sinds kort Bert, door jullie collegialiteit heb ik de afgelopen jaren het Leidse ROM-onderzoek kunnen promoten in binnen- en buitenland; er was altijd waarneming voor de klinische taken als ik op pad was.

Uitdrukkelijk wil ik de ROM-testverpleegkundigen uit het LUMC en Rivierduinen bedanken. Zij vormen de 'ruggengraat' van de grootschalige dataverzameling, en ik ben ervan overtuigd dat het belang van hun taak ernstig onderschat wordt. Een aantal van hen heeft mij ook jarenlang geholpen met de uitvoering van de MASHBANK: Marieke, Lilo, Marjan, Thea, Doeska, Mirja, Annie, Renske, Ingrid. Dank voor het bruikbaar maken van de data in SPSS: Arjen, en later Marion. Ook de ROM-baselineclub van het eerste uur is belangrijk geweest bij het bruikbaar maken van de data en het ontwerpen van de eerste studies: Margien, Yanda, Esther, Rosalind, Nic en Erik.

Ik heb veel hulp gehad van stagiaires en semi-artsen: Sanne, Riemke, Gülçin, Chantal, Julia en Mark. Veel dank voor jullie enthousiasme en leergierigheid. Ook de leergierigheid en kritische vragen van de arts-assistenten hebben mij steeds geïnspireerd. Zonder de hulp van het secretariaat zou de MASHBANK er niet zijn, en zou dit boekje waarschijnlijk nooit afgekomen zijn. Dank Alice, Vivian, Lia, Saskia, Corine, Marja, Jeannette, Lennie en Bianca.

Ook niet onbelangrijk: mijn intervisiegroep: Carla, Josee, Mirjam, Agaath, Barbara, Ingrid, Zohra, Arjen, Klaas en Matthijs. Mijn 'oude' opleiders: Alex, Frank, Ton, Adriaan en Adriano; dank voor het stimuleren van het onconventionele denken en de 'academische lancering' zoals Ton het noemde. Verder, op de poli en de afdeling: Tineke, Ingrid, Auke, Barbara, Nadine, Ron, Gea, Mirjam, en alle anderen. Op het Gorlaeus en het CKCL: Roel, Nienke, Liane en Jolanda. En natuurlijk de (ex)-kantoortuinbewoners: Marie-José, Ellen, Margien, Klaas, Leonie, Arianne, Yvonne en Nienke.

Michiel, als bijna 'native speaker' las jij alle manuscripten door, waarvoor zeer veel dank. Ook super dat je 'invliegt' om paranimf te zijn. Collega Janssen, onze vergaderingen in de avonduren waren altijd zeer verhelderend. Ik waardeer het dat je mijn vasculair geneeskundige paranimf bent! Hein en Richard, top dat jullie mij uit de brand hielpen met de lay-out toen ik met een gebroken sleutelbeen en een hersenschudding op de bank zat. Mijn ouders Ton en Marianne, mijn oma, Frans en Herma, dank voor al jullie steun!

Tenslotte, Linda, Lara, Julie en Emma, jullie zijn mijn alles. Dank dat jullie mij jarenlang vergroeid met mijn laptop in huis hebben gedoogd.

