CHAPTER 1

Introduction
Preface

Although major depressive disorder (MDD) is one of the leading causes of the global burden of disease, not all patients respond favorably to current treatments and scientific breakthroughs into its causes are not foreseeable. One problem contributing to this state of affairs is that patients with MDD are highly dissimilar in relevant respects, such as in clinical presentation, course, genetics and underlying pathophysiological mechanisms. It has led to doubt on the usefulness of the current classification: perhaps the definition of MDD captures a too diverse category of patients to find underlying causes or effective treatments. In order to create more homogenous subgroups, a variety of clinical depression subtypes have been suggested without curing the problem hitherto. This is the starting point of this thesis; the main question concerns whether we can identify clinically relevant, data-driven subtypes of depression that represent more homogenous subgroups in order to improve clinical decision-making and treatment assignments.

Broadly we focus on three questions: what sort of categories are we looking for, what methods can we use to identify these categories and what can we learn from the available data? In other words, this thesis discusses theoretical, methodological and empirical issues related to the search of better subtypes, and thus research and treatment, of MDD. In the following, we will elaborate the background for our search of data-driven subtypes of MDD, after which we will specify our empirical, methodological and theoretical approaches to investigate such subtypes.

Background

The burden of MDD

Major depression is currently one of the biggest contributors to the global burden of disease and is expected to be the leading contributor worldwide in 2020 [1-3]. There are several reasons for this high burden of disease, i.e. the high numbers of years lost to disability. First, MDD is a highly prevalent disease with lifetime prevalence estimates of 14.6% in the ten high-income and 11.1% in the eight low- to middle-income countries [4]. Second, the number of years with depressive episodes is relatively high, as MDD tends to start at a young age (median age of onset 29-34) [5] and has a recurrent course for most patients [6]. Third, MDD is a highly disabling disorder with substantial difficulties in role functioning, and is associated with low marital quality, low work performance and low earnings [7, 8]. Patients with MDD are more severely disabled than patients with physical disorders such as chronic pain conditions, cancer, and heart disease because MDD affects social and personal role functioning more than those physical disorders [9]. The high loss of health also leads to high economic costs: MDD is estimated to be the most costly brain disorder in Europe [10]. Finally, the personal suffering from MDD can be so severe that patients consider depression worse than dying [11]. Among all mental illnesses, major depressive disorder is one of the strongest risk factors for suicide attempts and completed suicide [12, 13]. Unfortunately, for many patients treatments are insufficiently effective [14-17]. One of the factors contributing to this unfortunate state of affairs is the wide diversity of patients with MDD [18].
The heterogeneity of MDD

Patients with MDD, according to current classification systems such as the Diagnostic and Statistical Manual of Mental Disorders (DSM) [19] and International Classification of Disease (ICD) [20], are highly dissimilar, which challenges current research and clinical practice. By definition, patients with MDD share a certain degree of disability and a syndrome including feeling depressed or anhedonia present for at least 2 weeks (DSM-5 [19], Box 1). Since every individual is unique, patients will always differ to some extent. However, patients with MDD are heterogeneous in several clinically relevant aspects.

First, response to treatment differs among patients. Effect sizes of antidepressant drugs, psychotherapy, and exercise are estimated in the range of 0.2-0.3 (Cohen's $d$) [14-17], meaning that about 60% of the patients treated with the active treatment have a better outcome than the average person in the placebo treatment group [21]. Thus, for some patients standard treatments are effective, but for a substantial proportion of patients these are not more effective than placebo, while they may have harmful side effects. Higher symptom scores at baseline predict a better treatment response to antidepressants [15, 22-24], but apart from that few clues have been found to match patients with effective treatments [25, 26].

Second, the course of MDD is highly variable. First, the duration of episodes varies widely, with approximately 50% of patients recovering within three months, while 20% still suffers from MDD after two years [27]. Some patients have such chronic episodes that they do not recover at all: a prospective cohort study, following patients with a first onset of depression syndrome during 23 years, showed that 15% of patients the disorder did not remit: patients never get out of their first episode [28]. Second, in recovered patients, the rate of recurrence is highly variable, with considerable numbers of patients having an episode of depression just once in their lives as opposed to patients with many recurrences [28-30]. Predictors tested so far perform poorly in distinguishing the patients with a long or recurrent course from those with a more beneficial course, and general risk assessment instruments are lacking [6, 31, 32]. The combination of the diversity and the unpredictability of the course of MDD makes up a major clinical problem, as it leads to difficulties in treatment decisions on the long term, with overtreatment of some patients, and undertreatment of others [33].

Third, despite extensive research, the etiology of MDD remains largely unknown and current studies mainly suggest a lot of variation between patients [34]. A broad range of risk factors have been investigated as possible causes of MDD, such as genetic abnormalities, mono-amine deficiency, disturbance of the hypothalamus-pituitary-axis, traumatic life experiences, low social support and education, and personality characteristics such as neuroticism and low self-esteem [34-37]. Although all these risk factors were found to be related to the development of MDD, correlations are weak in general [36, 37] (Figure 1). This could be the consequence of averaging an etiological diverse group of people: for some a risk factor might be strongly related to the development of MDD, but for others it might not be the case and thus the average result is a weak correlation [38]. Indeed, these findings have led to doubt on the biological homogeneity of MDD: it is unlikely that the current classification coincides with specific biological abnormalities but instead is capturing a group of people with different biological disturbances [39, 40]. The etiological heterogeneity is unfortunate as it hampers finding effective ways to interfere in the undesirable processes leading to MDD [41, 42].
Box 1. Diagnostic Criteria for Major Depressive Disorder (DSM) [19]

A. Five (or more) of the following symptoms have been present during the same 2-week period and represent a change from previous functioning; at least one of the symptoms is either (1) depressed mood or (2) loss of interest or pleasure.

1. Depressed mood most of the day, nearly every day, as indicated by either subjective report (e.g., feels sad, empty, hopeless) or observation made by others (e.g., appears tearful).
2. Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day (as indicated by either subjective account or observation).
3. Significant weight loss when not dieting or weight gain (e.g., a change of more than 5% of body weight in a month), or decrease or increase in appetite nearly every day.
4. Insomnia or hypersomnia nearly every day.
5. Psychomotor agitation or retardation nearly every day (observable by others, not merely subjective feelings of restlessness or being slowed down).
6. Fatigue or loss of energy nearly every day.
7. Feelings of worthlessness or excessive or inappropriate guilt (which may be delusional) nearly every day (not merely self-reproach or guilt about being sick).
8. Diminished ability to think or concentrate, or indecisiveness, nearly every day (either by subjective account or as observed by others).
9. Recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide.

B. The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.

C. The episode is not attributable to the physiological effects of a substance or to another medical condition.

D. The occurrence of the major depressive episode is not better explained by schizoaffective disorder, schizophrenia, schizophreniform disorder, delusional disorder, or other specified and unspecified schizophrenia spectrum and psychotic disorders.

E. There has never been a manic episode or a hypomanic episode.

Lastly, the definition of MDD in itself allows considerable variation in symptomatic presentation. MDD is defined in terms of a set of depressive symptoms, present for a period of at least two weeks, which leads to significant distress or impairment in functioning (Box 1) [19]. For the diagnosis, at least five out of nine depressive symptoms including at least one out of the two core symptoms depressed mood and anhedonia has to be present. This definition implies that depressive patients can potentially suffer from 227 different symptom combinations. Thus, the diagnosis for MDD allows considerable variation in the number and sort of depressive symptoms. Although no studies so far have studied the actual occurrence of those 227 combinations, there are indications that depressive patients have different symptom profiles dependent on age of onset [43] or the type of life events preceding the onset [44]. Moreover, specific symptom profiles might have consequences for prognosis: somatic symptoms such as weight loss, insomnia and fatigue predict a worse prognosis for patients after myocardial infarction [45] and cognitive symptoms as sad mood and concentration problems lead to more overall impairment [46].
**Figure 1.** Path and correlation estimates of Best-Fitting Model for Prediction of an MDD episode in the last year in 1,942 female twins

Two-headed arrows present correlation coefficients, and one-headed arrows represent path coefficients or standardized partial regression coefficients. Latent variables – indexed by observed variables in a measurement model – are depicted in rectangles. All variables have residual variance not depicted in the figure. Colors represent different periods associated with the development of MDD (blue): childhood (green), early adolescence (yellow), late adolescence (orange), adulthood (pink), the last year (purple). Figure reprinted from “Toward a Comprehensive Developmental Model for Major Depression in Women”, Kendler et al. [36], with permission.

**Clinical subtypes of MDD**
The problem of heterogeneity motivates the search for MDD-subtypes with similar clinical characteristics. Ideally, subtypes would describe MDD-patients that are similar with respect to multiple outcomes such as in terms of their disease persistence, severity and treatment reaction and who are different from patients with another subtype [47]. So far, those subtypes have not been identified [48]. Traditionally, subtypes have been proposed based on clinical pattern...
recognition and ordering, such as the subtypes described in the Research Diagnostic Criteria, the fore-runner of the DSM-III [49, 50]. Also the other disease categories in the DSM, such the definition of MDD itself, have been developed in this way, i.e. by psychiatrists in a process of expert meetings, study of the literature and consensus (sometimes termed the “bogsat” method, i.e. “bunch of guys sitting around a table”) [51].

Currently, at least 15 subtypes are described in the literature, based on symptoms, etiology, time-of-onset, gender or treatment-response (Figure 2) [48]. The value of these distinctions is contested. Although traditional subtypes are often associated with certain clinical characteristics (e.g. higher numbers of symptoms in case of melancholic depression) [52, 53], they do not clearly separate patients with a different etiology, course of illness, or treatment reaction [48, 54]. For instance, patients with two of the subtype distinctions described in the DSM5, melancholic and atypical depression, are remarkably similar with respect to family history, course pattern or demographic variables [53], which limits the epistemic value of these distinctions. Furthermore, current subtype distinctions are defined in different ways: they are either based on symptoms, supposed causes, gender, etcetera. It is unclear, however, what kind of variables (or combinations thereof) should be chosen to lead to successful subtype distinctions.

**Figure 2.** Current proposals for clinical subtypes of MDD

<table>
<thead>
<tr>
<th>Aetiologically-based:</th>
<th>Time of onset-based:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Adjustment disorder</td>
<td>- Early trauma depression</td>
</tr>
<tr>
<td>- Reproductive depression</td>
<td>- Perinatal depression</td>
</tr>
<tr>
<td>- Organic depression</td>
<td>- Drug induced depression</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Symptom-based:</th>
<th>Gender-based:</th>
<th>Treatment response-based:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Melancholia</td>
<td>- Female depression</td>
<td>- Treatment resistant depression</td>
</tr>
<tr>
<td>- Psychotic depression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Atypical depression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Anxious depression</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

15 frequently proposed clinical subtypes of depression in 754 reviews published between 2000-2011, categorized according to different starting points for subtyping. Arrows indicate that subtypes are not mutually exclusive: patients can satisfy more than one subtype. Figure reprinted from “Meta-review of depressive subtyping models”, Baumeister and Parker [48], with permission from Elsevier.
This thesis

Aim of this thesis: empirical subtypes of depression
In short, the starting-point of this thesis is the following: MDD-patients suffer from a high burden of disease, but are dissimilar in many respects, which complicates treatment decisions and research. Clinical subtypes of depression have been introduced as a means to reduce the heterogeneity, but these have not resulted in patient groups with a more similar course of illness, etiology or treatment response. However, what is not yet well understood is whether new empirical methods could lead to homogeneous subtypes of MDD with a higher clinical value. The aim of this thesis is to investigate the existence of such data-driven subtypes of depression.

Data-driven subtypes of depression are informed by similarity patterns in data identified by statistical analyses. As mentioned before, definitions for subtypes have traditionally been based on pattern recognition by clinicians, resulting in categories without sufficient predictive value in clinical practice. As an alternative, data-driven strategies have been adopted in the search of MDD subtypes. In this case, statistical methods are used to discern similarity patterns in data of MDD-patients, which are used to underpin potential subtype distinctions. The advantage of using statistical techniques for pattern recognition could be that different and unknown patterns and trends emerge from the data providing insight in relevant subgroups, which have not been identified by clinical pattern recognition [55, 56].

The very possibility of identifying data-driven subtypes is fairly new and still expanding – dependent on new statistical methods and increased computation speed in large datasets [56] – which makes the approach in this thesis innovative. The novelty of this research field also implies that it is unlikely that we will discover definite, clinically useful, data-driven subtypes of depression in this thesis. Rather, we intend to formulate starting-points relevant to the search of data-driven subtypes of MDD by focusing on three related questions covering theoretical, methodological and empirical aspects. First, what empirical evidence is currently available for data-driven subtypes of depression? Second, if we aim for data-driven subtypes of depression, what sort of statistical methods can we use to identify them? The last question is more theoretical and concerns the nature of disease classification. If we search for data-driven subtypes of depression, what sort of categories are we looking for? Below we will further introduce these three questions.

Empirical evidence for data-driven subtypes
The first studies concerning data-driven subtypes of MDD have been published in the 1960’s and 1970’s, of which most applied factor analyses, and some cluster analyses, to identify patterns in data of depressed patients (e.g. [57-59]). Since the introduction of the current definition of MDD in 1980 [60], studies have occasionally used data-driven methods but only recently the field started to grow (e.g., [61, 62]). Given the limited value of clinical subtypes of depression, our first question concerned the available evidence for data-driven subtypes of depression. Have stable patterns been observed in the data up to now, and if so, what patterns? Secondly, if we study data-driven subtypes of MDD in new datasets, what patterns emerge from these analyses? Are they comparable to previous findings? Which characteristics discriminate between patients with different subtypes? In short, what can we learn from available data about subtypes of depression?

Methods to discover data-driven subtypes
In search of data-driven subtypes of depression, one inevitably encounters methodological and statistical questions. There is no free lunch in statistics: there is not one best statistical method,
but different methods might be suited for different problems [63]. In addition, different models might give different pictures of the relation between the predictors and the outcome [64]. Hence the second question of this thesis is of a methodological nature and concerns an investigation of statistical methods that could be applied in search of MDD-subtypes.

One major choice in statistical methods used in search of MDD-subtypes is between supervised or unsupervised methods [56, 63]. In case of supervised methods, one investigates the relation between a set of predictors (e.g., a variety of initial depressive symptoms) and an outcome variable (e.g., hospitalization for MDD). The aim of such methods is to accurately predict the outcome for future patients (prediction) or to understand better the relationship between the predictors and the outcome (inference) [63]. One question for supervised learning methods such as linear regression, linear discriminant analyses and regression trees would be: which depressive symptoms predict hospitalization for MDD and how strong are their effects? In contrast, unsupervised learning methods are used when there is no outcome variable but only a set of characteristics such as depressive symptoms. In this case, the goal of statistics is to describe how the data are organized or clustered in order to understand the relationships between the predictors. For instance, there are statistical methods that measure the degree of similarity or dissimilarity between subjects based on the initial depressive symptoms, and accordingly classify patients into subgroups with similar depressive symptoms. The hope is that those homogeneous subgroups will coincide with certain other clinically relevant characteristics such as etiology or reaction to treatment [61, 63]. Examples of unsupervised methods are latent variable analyses, such as latent class analyses and exploratory factor analyses. This thesis investigates variants of both supervised and unsupervised methods, and applies new supervised learning techniques in the search of subtypes of MDD.

Comorbidity and its implications for the nature of disease classification

In search of subtypes of depression, one inevitably encounters questions regarding the nature of such classifications. What sort of structures are we looking for? And how should we interpret the definitions of psychiatric disorders currently described in the DSM? These questions come clearly to the fore in the discussion on psychiatric comorbidity. Comorbidity is rather the rule than the exception in psychiatry, with 35% to 45% of patients reporting two or more disorders in the course of one year [65-67]. Comorbidity occurs even more frequently in patients with MDD: about three quarters of depressive patients suffer from additional psychiatric disorders in the course of one year, mainly anxiety disorders, alcohol dependence and ADHD [67]. The high rates of comorbidity have led to an extensive theoretical debate on the nature of psychiatric disorders. Why do we have so much comorbidity in psychiatry? What do those high rates of comorbidity tell us about psychiatric disease classifications? This thesis studies psychiatric comorbidity in order to clarify the kind of structures we are looking for.

General relevance

With this thesis we aim to provide some novel perspectives on subtypes of MDD, informed by insights from psychiatry, statistics, and philosophy of science, in order to resolve part of the heterogeneity of MDD. However, the relevance of this thesis might go beyond research on MDD only. The problem of heterogeneity applies to more psychiatric disorders, such as to schizophrenia, attention deficit hyperactivity disorder, and borderline personality disorder [68-70]. Some of the insights concerning data-driven subtypes of MDD might benefit research trying to provide solutions and subclasses of other heterogeneous psychiatric disorders.
This thesis also connects to a broader debate about the validity of classifications in psychiatry, which intensified during the writing of this thesis with the development and introduction of the fifth edition of the DSM (DSM-5, 2013) [19]. Some have heavily criticized certain classification decisions [71], and others have even doubted the validity and reality of psychiatric disorders in general [72, 73]. We hope that this thesis contributes to a better understanding of psychiatric disorders and to a more constructive debate about psychiatric disease classifications.

**Thesis outline**

The first part of this thesis (chapter 2-6) focuses on the methodological and empirical questions described above, viz. what methods can we use to identify data-driven subtypes of depression and what evidence do we get from the data available? First, we performed a systematic review of current evidence available for data-driven subtypes of depression identified by unsupervised, latent variable methods (chapter 2). Second, we aimed to identify subtypes of depression that directly predict course of illness based on initial symptoms in a large general population sample (chapter 3). This study served as a proof-of-concept as it tested whether it was possible to relate subtypes directly to a clinical relevant outcome using a novel combination of supervised statistical learning techniques. In a follow-up study, we tried to improve the performance of these identified data-driven subtypes by expanding the analyses with information on comorbidity patterns (chapter 4). Fourth, we investigated possible additional predictors related to course of illness by studying a rich set of risk factors for recurrence of depression in a prospective study of females with MDD (chapter 5). Fifth, we investigated the value of statistical interactions for prediction models in medicine (chapter 6). For this aim, we studied a different patient population, viz. patients with myocardial infarction, for whom good prediction models are already available [74, 75].

The second part of the thesis (chapter 7-9) concerns the more theoretical aspects related to data-driven subtypes of MDD. We studied psychiatric comorbidity in order to clarify the nature of psychiatric disorders. We start this section by an analysis of one part of the discussion on comorbidity – is comorbidity a problem or a validator? – and discuss how this discussion relates to the absence of causal disease models in psychiatry (chapter 7). Next, we investigate the influence of classification choices on rates of comorbidity. Certain classification choices, such as diagnostic thresholds, are suspected to have substantial influence on psychiatric comorbidity. In a large sample from the general population, we studied to what extent different classification choices influence comorbidity between MDD and GAD (chapter 8). Then, we put these empirical findings in a theoretical perspective and discuss their consequences for the interpretation of comorbidity and psychiatric disorders in general, by relating the debate on psychiatric comorbidity to other debates in the philosophy of science (chapter 9). In the last chapter of this thesis we will discuss our findings in a broader context and come back to the questions described above (chapter 10).
References


Introduction


