Exploring paths forward for mood disorder research:

Embracing heterogeneity, complexity and idiography

Kaat Hebbrecht

Exploring paths forward for mood disorder research:

Embracing heterogeneity, complexity and idiography

Proefschrift voorgelegd tot het behalen van
de graad van doctor in de Medische Wetenschappen
aan de Universiteit Antwerpen
en te verdedigen door:

Kaat Hebbrecht

Promotores:	Prof. dr. B. Sabbe
	Prof. dr. E.J. Giltay
Begeleider:	Prof. dr. M. Morrens
Leden van de interne doctoraatscommissie:	Prof. Dr. Paul Van Royen
	Prof. Dr. Geert Dom
Leden van de externe doctoraatscommissie:	Prof. Dr. Stephan Claes
	Prof. Dr. Robert Schoevers

Antwerpen, 2022

Cover:
Silke Reyntjens (Sire)

Layout en printer:
Proefschriftenmaken.nl

ISBN/EAN:
978-94-6469-119-1

Table of contents

Preface

The major mood disorders, that include major depressive disorders (MDD) and bipolar disorders (BD), are among the most critical health-related issues of our time. Yet, despite many research efforts, progress in understanding the etiopathogenetic mechanisms underlying the onset and course of mood disorders is slow. Furthermore, treatments are often ineffective, especially those for depressed states, where one third of 'real-world' patients do not respond to four courses of antidepressant treatment. One of the contributing factors to the slow scientific progress could be the diagnostic constructs, and their inherent limitations, that are currently used in mood disorder research. Another factor relates to the inherent complexity of mood disorders, and psychopathology in general, as they result from complex interactions between psychological, neurobiological and social factors over time. Consequently, the course and underlying etiopathogenetic mechanisms of mood disorders are highly individual-specific. In this doctoral thesis, we explore paths forward for mood disorder research that may help increase our insight into the course and underlying etiopathogenetic mechanisms of mood disorders.

CHAPTER 1

General introduction

In the first chapter of this dissertation, a general background of mood disorders and the current diagnostic classification criteria are provided, after which the challenges that mood disorder research is currently facing are highlighted. Next, three paths forward for mood disorder research are presented that may provide deeper insight into the underlying etiopathogenesis and course of mood disorders. These include: 1) a symptom-oriented approach, 2) the investigation of the complex interplay of symptoms over time and, 3) individual-level (i.e., idiographic) analyses. This chapter concludes with a description of the general aim of the thesis, the studied samples and design of the studies used in this thesis and a general outline of the thesis.

1.1 Mood disorders

Mood disorders constitute a group of psychiatric syndromes in which a disturbance in mood is the most prominent clinical feature. Major depressive disorder (MDD) is characterized by a prominent depressed mood, whereas in bipolar disorder (BD) unusual fluctuations in mood states are seen, ranging from episodes of elation ((hypo)mania) to episodes of depression [1]. Cognitive dysfunction is another central feature of mood disorders that is closely related to the functional impairment that these disorders frequently cause [2–5]. The world-wide 12-month prevalence of mood disorders is estimated at 4.4 %, which is equivalent to more than 300 million people of the total world population [6].

Patients with MDD or BD generally show a substantial and often long-standing impairment in work, social and family domains and subjective wellbeing [6,7]. Their family members are also confronted with a high level of suffering and some even develop a depressive disorder themselves [8]. According to the World Health Organization (WHO), mood disorders are the leading cause of disability worldwide (with MDD accounting for 7.5% and BD for 1.3 % of all years lived with disability) [6] and they are associated with a higher risk of other psychiatric disorders (e.g., anxiety and substance use disorders) [9,10] and critical somatic disorders such as cardiovascular diseases [11–14]. Together with an increased risk of suicide, the latter conditions contribute to the raised mortality rate that is observed in people coping with mood disorders compared to healthy peers [15].

The economic burden associated with mood disorders is similarly disconcerting. The largest component derives from indirect costs such as work disability and a slightly smaller proportion from direct medical costs (in- and outpatient care, medication and emergency costs) [16]. The total costs for mental health care in Belgium were estimated at 5% of the gross domestic product (GDP, or 'bruto binnenlands product' BBP) in 2015, making Belgium (together with The Netherlands and Finland) one of the unfortunate front runners in Europe [17].

Overall, the human, social and economic facts described above make understanding of mood disorders one of today's most compelling health-related challenges that deserve our thorough research.

1.2 Diagnostic Classification of Mood disorders

A reliable diagnosis of mood disorders is of importance for several reasons: to establish a common language in the clinical diagnosis of patients, to inform treatment prognosis and to gather quantifiable public health information (e.g., prevalence, morbidity and mortality indices) [18,19]. Furthermore, operationalized diagnostic constructs are essential for research into the underlying etiopathogenetic mechanisms of mood disorders. The DSM-5 (Diagnostic and Statistical Manual of Mental Disorders – fifth edition) [1], next to the ICD-11 (International Classification of Diseases and Related Health Problems) [20], is the most widely used classification manual for psychiatric diagnoses. The DSM has provided an international agreed diagnostic classification system, enabling clinicians and researchers to describe and communicate more reliably about psychiatric diagnoses. Yet, the use of DSM diagnoses for research purposes also has its pitfalls, which will be discussed further on in this thesis.

The DSM-5 classifies MDD and BD as separate disorders distinguishable by a history of (hypo)mania. Bipolar disorders (BD) are categorized between the schizophrenia spectrum disorders and MDD because of the evidence suggesting considerable overlap in the symptoms, family history and genetics of BD with the two other disorders [21]. A distinction is made between type-I and type-II BD. The former is defined by the occurrence of manic episodes and, non-obligatory, major depressive episodes (MDE); whereas the latter is characterized by the presence of hypomanic episodes and MDE. Of note is that the manual makes no distinction between a depression occurring within MDD or BD, both referred to as MDE. The DSM-5 criteria for MDE and (hypo)mania are presented in the text boxes below.

DSM-5 Criteria for major depressive episode (MDE)

A. ≥5 symptoms during the same two-week period that are a change from previous functioning; depressed mood and/or loss of interest/pleasure must be present; exclude symptoms clearly attributable to another medical condition

1. Depressed mood
2. Loss of interest or pleasure
3. Significant weight loss or weight gain, or decrease or increase in appetite
4. Insomnia or hypersomnia
5. Psychomotor agitation or retardation
6. Fatigue or loss of energy
7. Feelings of worthlessness or excessive or inappropriate guilt
8. Diminished ability to think or concentrate or indecisiveness
9. Recurrent thoughts of death or recurrent suicidal ideation or attempt.

B. The depressive symptoms need to be accompanied by clinically significant distress or functional impairment
C. The symptoms are not attributable to the physiological effects of a substance or another medical condition.

DSM-5 Criteria for hypomania

A. A distinct period of abnormally and persistently elevated, expansive, or irritable mood and abnormally and persistently increased activity or energy, lasting at least 4 consecutive days and present most of the day, nearly every day.
B. During the period of mood disturbance and increased energy or activity, three (or more) of the following symptoms have persisted (four if the mood is only irritable), represent a noticeable change from usual behavior, and have been present to a significant degree:

1. Inflated self-esteem or grandiosity
2. Decreased need for sleep (e.g., feels rested after only 3 hours of sleep)
3. More talkative than usual or pressure to keep talking
4. Flight of ideas or subjective experience that thoughts are racing
5. Distractibility (i.e. attention too easily drawn to unimportant or irrelevant external stimuli), as reported or observed
6. Increase in goal-directed activity (at work, at school, or sexually) or psychomotor agitation
7. Excessive involvement in pleasurable activities that have a high potential for painful consequences (e.g., engaging in unrestrained buying sprees, sexual indiscretions or foolish business investments)

C. The episode is associated with an unequivocal change in functioning that is uncharacteristic of the person when not symptomatic
D. The disturbance in mood and the change in functioning are observable by others
E. The mood disturbance is not severe enough to cause marked impairment in social or occupational functioning, or to necessitate hospitalization, and there are no psychotic features
F. The symptoms are not due to direct physiological effects of a substance (e.g., a drug of abuse, a medication or other treatment) or a general medical condition (e.g., hyperthyroidism)

DSM-5 Criteria for mania

A. A distinct period of abnormally and persistently elevated, expansive, or irritable mood and abnormally and persistently goal-directed behavior or energy, lasting at least 1 week and present most of the day, nearly every day (or any duration if hospitalization is necessary).

B. During the period of mood disturbance and increased energy or activity, three (or more) of the following symptoms have persisted (four if the mood is only irritable) are present to a significant degree and represent a noticeable change from usual behavior: *see criteria 1 to 7 in box DSM-5 criteria for hypomania*

C. The mood disturbance is sufficiently severe to cause marked impairment in social or occupational functioning or to necessitate hospitalization to prevent harm to self or others, or there are psychotic features.

D. The episode is not attributable to the direct physiological effects of a substance (e.g.,, a drug of abuse, a medication, or other treatment) or another medical condition.

1.3 Mood disorder research

1.3.1 Previous findings

Research in the last two decades has focused on the etiopathogenetic mechanisms that underly the onset and course of mood disorders, addressing biological, sociological, and psychological variables. Although important knowledge has been accrued, several basic questions remain. As for the biological findings, genome-wide association studies have failed to generate robustly replicated results on genetic loci associated with MDD or BD [22–24] and no consistent biomarkers could be identified that were solidly and specifically associated with MDD or BD [25]. Furthermore, accurate prediction of the onset and course of mood disorders has proven to be challenging [26,27].

In the following section of this thesis, we will discuss three challenges for mood disorder research that may have contributed to this apparent lack of scientific progress. First, we address the dominant methods used to diagnose and conceptualize psychopathology (i.e., the study of MDD and BD as syndrome-based disease entities). Although relevant for providing a common language for clinical purposes, these methods rely on assumptions that are inherently problematic and hereby may have slowed down research progress. Second, many of the current research practices we employ to investigate the onset or course of mood disorders do insufficient justice to the dynamic nature of symptoms. Third, even though it is common knowledge that symptoms and etiopathogenetic mechanisms of mood disorders can vary highly among and within patients, individual-specific analyses have insufficiently been performed in research endeavors.

1.3.2 Challenges in mood disorder research

Lack of conceptual clarity on a research framework for mood disorders

Over the past decades, research in mood disorders has been mainly conducted at the level of syndromes, i.e. clusters of symptoms that co-occur [28]. Specifically, study populations were based on DSM (or ICD) criteria and studies were aimed at elucidating risk factors, biomarkers or treatment response of the syndrome MDD (or BD) [29,30]. These practices start from the basic assumption that syndromes (such as MDD or BD) are *discrete disease entities* for which an underlying causal *essence* can be found [31,32]. Thus, the dominant explanation for the coincidence of symptoms was that they all stem from a shared causal origin (known as the *common cause approach*) [31,33]. However, these assumptions run up against important boundaries.

The proposition that syndromes such as MDD and BD are *distinct disease entities* is challenged by several research findings. First, the quest for the common etiopathogenetic pathway of syndromes such as MDD (or BD) has yielded limited insights. As alluded to above, genome-wide association studies have failed to identify robustly replicated results on genetic loci associated with MDD or BD [22–24] and no consistent biomarkers were solidly associated with either MDD or BD [25]. Growing evidence rather points to the existence of transdiagnostic risk factors [34,35] and shared biomarkers across psychiatric disorders [36]. Also, symptom heterogeneity within DSM constructs is high, which is partly inherent to the DSM criteria themselves. For example, there are 227 possible ways to reach a diagnosis of MDE using the DSM-5 criteria. Fried et al. [37] demonstrated that this *possible* variation in symptom profile also translates in *true* variability by identifying up to 1030 unique symptom profiles in depressed outpatients in the STAR*D [38,39] sample. Moreover, there is a high rate of co-morbidity among the DSM constructs [40,41]. For example, MDD is highly comorbid with anxiety disorders and substance use disorders [42–44]. Finally, the DSM's diagnostic boundaries seem to be somewhat arbitrary as mounting evidence supports the dimensional character of psychopathology with stages of mild symptoms at the one extreme to full-blown psychiatric disorder at the other [45,46]. So, although few would question the value of standardized diagnostic criteria as a clinical tool, the strict adherence to DSM-defined syndromes may have hindered research progress in the underlying mechanisms of mood disorders. This has also been acknowledged by the editors of the "Research Agenda to the DSM-5" [41].

> *"The over-reification of the DSM categories has led to a form of closed-mindedness on the part of researchers and funding sources. For example, researchers involved in new drug development tend to focus their efforts on treatment of DSM-IV-defined categories, despite widespread evidence that pharmacologic treatments tend to be effective in treating a relatively wide range of DSM disorders. Furthermore, the erroneous notion that the DSM categories can double as phenotypes may be partly responsible for the lack of success in discovering robust genetic markers.[p 34]"*

The common-cause theory – all symptoms have a shared causal origin - further includes important assumptions on how we perceive symptoms. Symptoms are regarded as passive consequences of the underlying disease construct and accordingly treated as equal, interchangeable and independent factors of the this underlying construct (e.g., MDD or BD) [47]. In research, this is reflected by the pervasive use of sum scores: item scores of individual symptoms are generally added up into a sum-score of depression rating scales, where this sum score is interpreted as a reflection of severity or an indicator of treatment response of the underlying 'unitary construct' [30,48]. However, the equivalence of depression symptoms has repeatedly been called into question. Specifically, more and more studies show that individual symptoms of depression have different risk factors [49,50], psychosocial impact [51] and course trajectories [52]. Moreover, individual symptoms can show differential effects to antidepressant treatments [53,54]. Further, not all depressive symptoms are correlated with each other. Some are more strongly connected with each other and more important for the persistence (or recovery) of the underlying disorder [55].

Taken together, there is a need for alternative conceptual frameworks for mood disorder research that transcends the limitations of the categorical classification systems and that goes beyond the use of aggregated sum scores. One alternative research framework, introduced by the US National Institute of Mental Health, is the Research Domain Criteria (RDoC) project which proposed to study dimensions of observable behavior and neurobiological variables that correlate with biological (e.g., genetic, neuroimaging and neuropsychological) factors, irrespective of diagnostic boundaries [56].

The dynamic nature of psychopathological symptoms

A second challenge for mood disorder research is to incorporate the dynamic nature of symptoms in modelling approaches. Mood disorders have predominantly been studied as categorical entities and modelling techniques were largely static [27,29,57]. Specifically, the course of MDD and BD is mainly investigated using static outcomes (e.g., response or remission rates) that are based on the change in mean or on threshold scores of symptom rating scales and hereby assuming a gradual linear change over time [27]. In real life, however, symptoms show a dynamic, non-linear behaviour over time and interact

with other symptoms such that the onset of one symptom may influence the onset and/or persistence of others [58]. Clearly, research into the course and prediction of mood disorders is then likely to benefit from studying the patterns of symptom changes and symptom-symptom relationships over time as these could provide a more nuanced insight into the symptom processes underlying the development of psychopathology.

Group-to-individual generalizability of research findings

A third challenge in psychiatry research is the high level of interindividual variability. Individual patients differ considerably in terms of their symptom profile and symptom patterns over time [37,48] and the underlying etiopathogenetic mechanisms may even be unique for individual patients and this may hold true for various disorders [37,59,60]. Nevertheless, most psychopathology research employs group-level (i.e., nomothetic) analyses. Although they are important to make general predictions for the studied population, these group-level results do not always translate to the level of the individual clinical patient [61,62]. Complementing group-level with individual-level analyses could increase our knowledge of patient-specific symptom mechanisms that drive the development and persistence of psychopathology, knowledge that may bear great relevance for treating clinicians.

1.3.3 Potential ways forward for mood disorder research

Shifting from a syndrome-oriented to a symptom-oriented approach

An alternative approach for research is to cut across current diagnostic boundaries and shift our attention from whole-syndrome analyses to smaller units of analysis [63]. One possible way is to focus on individual symptoms and to study the distinct patterns of behavior of symptoms and their effects on other symptoms [30,37]. Network analysis is a promising framework to study associations between symptoms and it provides an alternative conceptualization of psychopathology compared to the syndrome-oriented approaches. Specifically, correlations between symptoms are not explained by a common cause (e.g., MDD or BD) but psychopathology is rather conceptualized as a network of causally connected symptoms. The connection between specific symptoms may reflect underlying biological processes (e.g., insomnia leading to fatigue that subsequently leads to concentration problems) and/or psychological processes (e.g., negative thinking leading to hopelessness). In the network framework, symptoms are not seen as passive consequences of an underlying disorder but rather as autonomous and adaptive entities that can exert an effect on other symptoms (Figure 1). Furthermore, from a network perspective, not only symptoms and their interrelationship can be studied but also the structure of the overall network (e.g., strongly or weakly connected) [47,64,65].

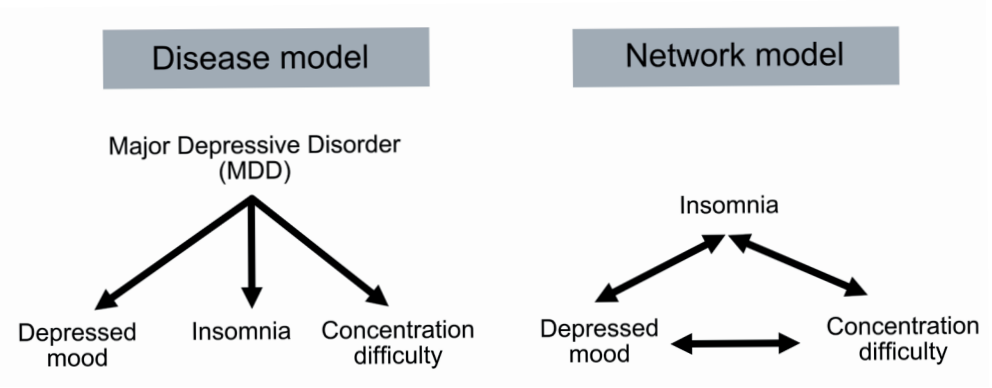


Figure 1. Disease model versus network model of psychopathology

The premise here is that examining individual depression symptoms and their complex interactions will promote our understanding of the heterogeneity in the symptoms defining mood disorders [37,48]. Further, this approach could provide a more plausible explanation of co-morbidity (i.e., interactions between symptoms) than the syndrome-oriented approach (i.e. two distinct disorders) [29].

Moving from a static to a dynamic approach

There is a growing interest in research that addresses psychopathology as a complex dynamic system. In this dynamic framework, psychopathology is assumed to result from direct relations between symptoms that belong to the same system. Specifically, symptom-symptom interactions are thought to give rise to emergent properties such as depression or other mental disorders (or the resolution thereof) and these outcomes cannot be predicted from one single symptom alone. Further, the system of symptom-symptom relations can evolve within an individual over time [27,66]. These symptom-symptom interactions may occur on different time-scales and the behavior of the system is non-linear, making it hard to predict over time.

Although this complex dynamic system approach to psychopathology is not a new concept [67–69] , its implementation in research was limited by the demand for intensive longitudinal data and the lack of suitable statistical models for the estimation of dynamic change. Recent advances in technology (e.g., electronic data collection, mobile devices) and methodology have greatly facilitated the study of dynamic (symptom) processes in psychopathology. Although the network framework has mainly been used to investigate cross-sectional relations between symptoms [70], recent network analytic studies also investigated relations between individual symptoms from one time point to another, using newly developed statistical methods, especially lagged multilevel vector autoregressive

(VAR) models [71]. This method captures the relationship between symptoms from one time point to the next and combines both individual-level as well as group-level dynamics into the model.

Investigating mood disorders as complex dynamic system provides a means to investigate which symptoms are more influential in a dynamic symptom network compared to other symptoms and thus could serve as optimal targets for treatment intervention. Furthermore, the investigation of the characteristics of the overall structure of the dynamic symptom network may inform prediction of MDE onset and recovery of MDE.

Combining group-level and individual-level approaches

There is a large heterogeneity among patients with mood disorders in terms of their symptom trajectories and (cognitive) side effects to treatment over time. Individual-specific analyses have seldom been performed in Mood Disorder research. However, they provide us with a unique opportunity to study individual variability in a wide range of variables. Individual-specific analyses could be employed to study the dynamic behavior of symptoms and dynamic symptoms relations *within* individual patients. This idiographic approach provides an interesting gateway to study how symptoms evolve at the level of the individual in daily clinical practice and hereby facilitate a road towards personalized research and treatment [59,61,72].

Another area for which individual-specific analyses could be particularly relevant is the study of cognitive dysfunction in mood disorders. Potentially highly disabling [2–5], the presence and degree of cognitive impairment differ widely among patients with mood disorders [73,74]. This also applies to the cognitive effects of antidepressant treatments such as some psychopharmacological agents or electroconvulsive therapy [75–77]. And yet, most mood disorder studies to date examine cognitive functioning as a homogeneous phenomenon, reporting mean (group-level) averages for cognitive test outcomes. By combining group-level and individual-level analysis methods, we will be able to account for the interindividual variability in cognitive functioning in mood disorders and thereby facilitate its clinical interpretation.

1.4 General aim of the thesis

The general aim of this thesis is to gain a better insight into the course and underlying etiopathogenetic mechanisms of mood disorders. We used both traditional as well as novel data analytic approaches to achieve this aim.

PART 1: Traditional approaches

The first part of this thesis is based on more traditional methodological approaches (i.e., syndrome-oriented, static outcome measures, group-level analyses) to study the symptom profiles and course, and inflammatory biomarkers of MDD and/or BD. Specifically, we studied the clinical and biological variables in relation to syndrome-based entities such as MDD or BD. Next, using a group-based method, we looked at averages of outcome variables such as that of sum scores of depression rating scales and blood levels of inflammatory markers. Finally, we applied a static modelling approach to examine changes in mood or biological markers over time (as opposed to a dynamic approach which takes the dynamic interactions of the constituent elements of the disorder, such as symptoms or biological elements, into account).

Symptom profile and course of depressive episode in patients with MDD versus BD (Chapter 2)

The DSM-5 considers MDD and BD as two distinct disease entities, with evidence indicating the importance of a distinct treatment approach [78,79]. Early differentiation between MDD and BD is thus essential but poses a great challenge since the initial manifestation of BD is mostly a depressive mood state and no clinical pathognomonic features for a 'bipolar depression' have been found [80].

To investigate whether specific clinical characteristics of MDE could provide diagnostic clues for a differentiation between an underlying MDD or BD, in Chapter 2, we compared symptom profiles and the course of MDE in patients with MDD versus BD.

Tryptophan catabolites as a biomarker for BD (Chapter 3 & 4)

Accumulating evidence points to a dysregulation of the inflammatory system in mood disorders, such as elevated levels of peripheral pro-inflammatory cytokines [81,82] and activation of the microglia in the central nervous system [83,84]. The kynurenine pathway of tryptophan (TRP) degradation is one of the mechanisms through which inflammation leads to neurotoxicity and mood symptoms [85–87]. An overview of the kynurenine pathway is provided in Figure 2.

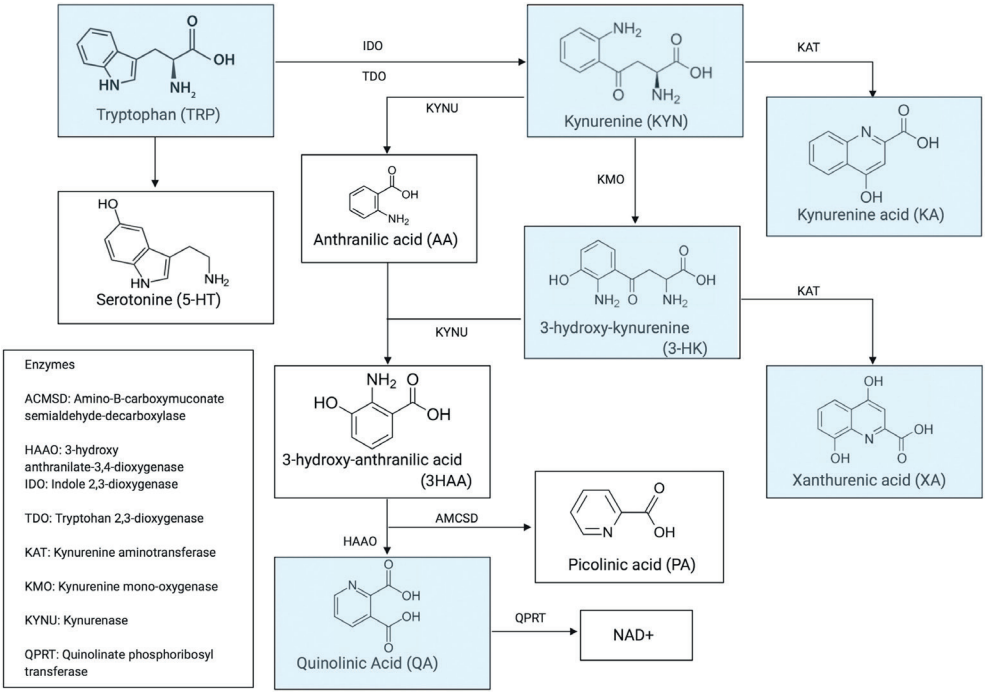


Figure 2. Kynurenine pathway

Although TRP is primarily known as a precursor of serotonin, about 95% of TRP is metabolized via the kynurenine pathway [87]. Two main hypotheses have been put forward to explain the possible role of the kynurenine pathway in the relationship between inflammation and neurotoxicity. The serotonin-depletion hypothesis states that, in response to inflammation or stress, TRP is mainly metabolized into kynurenine (KYN) at the expense of the formation of serotonin, where serotonin depletion is a well-established pathophysiological mechanism of mood disorders [88–90]. A second hypothesis poses that an imbalance of TRP catabolites (TRYCAT) occurring downstream from KYN in the brain plays a crucial role in neurotoxicity, independent of serotonin depletion. More specifically, a decrease in neuroprotective kynurenic acid (KA) catabolites (mainly in astrocytes) and an increase in supposedly neurotoxic 3-hydroxy kynurenine (3-HK) and quinolinic acid (QA; mainly in microglia) have been observed in patients with mood disorders. Kynurenic acid (KA) exerts its neuroprotective effect by NMDA-receptor antagonism and by counteracting the neurotoxic effect of QA, where QA and 3-HK are assumed to have a neurotoxic effect through oxidative stress and N-methyl-D-aspartate (NMDA) receptor agonism [86,87,91,92].

Meta-analyses of TRYCAT alterations in MDD revealed a consistent decrease in TRP, KYN and KA [93–95]. Studies in BD are, however, far fewer in number, leaving it unclear whether TRYCAT alterations in BD are different to those MDD. Furthermore, it is not known whether TRYCAT alterations in BD are mood-state specific. As several new studies comparing

TRYCAT levels in BD with those of healthy controls have been published in the last 5 years, we performed systematic review and meta-analysis of studies investigating TRYCAT alterations in cerebrospinal fluid and peripheral blood (Chapter 3). Additionally, we performed subgroup analyses for the different mood states in BD to examine whether differential TRYCAT alterations exist between (hypo)manic and bipolar depressed BD patients.

As TRYCAT alterations have been suggested to play a pathophysiological role in the cognitive deficits associated with BD [86], in Chapter 4, we aimed to assess whether patients with BD show different TRYCAT levels according to their mood state and whether these TRYCAT alterations are associated with cognitive dysfunction in BD.

PART 2: Idiographic approach to symptom dynamics (Chapter 5 & 6)

The second part of this thesis aims to provide a more fine-grained insight into the course of MDE by examining the dynamic symptom changes that lie at its basis.

Hence, in Chapter 5, we investigated the dynamic behavior of individual depression symptoms over time, based on repeated Hamilton Rating Scale for Depression (HRSD-17) measurements, using Dynamic Time Warp (DTW). These dynamic symptom changes were investigated both at the individual-level and the group-level. These analyses yielded undirected symptom dimensions (i.e., groups consisting of symptoms with similar dynamics over time).

Undirected DTW analyses do not allow to disentangle whether changes in a symptom temporally precede or follow changes in another symptom. Since this is a necessary (albeit not sufficient step) towards causal inference and ultimately clinical interventions, in Chapter 6, we investigated both undirected and directed depression symptom dynamics (which investigate the directionality of symptom change over time) based on repeated Beck Depression Inventory (BDI-II) data.

PART 3: Cognitive function in MDE patients treated with ECT (Chapter 7 & 8)

The third part of this thesis aims to provide a more nuanced insight into the cognitive effects of ECT in patients with MDE by combining group-level and individual-level analyses of cognitive function.

In Chapter 7, we investigated the short- and long-term effects of ECT on mean global cognition in patients with MDE. Group-level changes in (sub)scores of the Montreal Cognitive Assessment (MoCA) were investigated using linear mixed models (LMM) and individual changes were analyzed using Reliable Change Indices (RCI).

In the study included in Chapter 8, we assessed a wide range of cognitive domains and investigated their evolution over time both at the group-level (using LMM) and at the individual-level (using RCI). Subsequently, we investigated whether we could identify distinct classes of patients showing similar cognitive trajectories over time, using Latent Class Growth Analysis (LCGA).

Table 1. Characteristics of the data and analysis approaches used in this thesis

	Participants	Sample size	Chapter with aim	Data analysis approach	
Hercules	MDE	n = 224 MDD n = 52 BD	**Chapter 2:** To study the symptom profile and course of MDE in MDD versus BD	- Average means of HRSD-17 subitems in MDD and BD - Change in mean HRSD-17 sum-scores over time	PART 1
Case-control studies	BD (D/M/E)	n = 850	**Chapter 3:** To determine the association between TRYCAT alterations in CSF or peripheral blood in BD	Studying mean TRYCAT levels of syndrome-based groups (BD, BD-D, BD-M)	
Inflammation and Neuroprotection in BD	BD (D/M)	n = 67	**Chapter 4:** To examine 1) whether patients with BD show different TRYCAT alterations according to their mood state, and 2) whether TRYCAT alterations are associated with cognitive dysfunction in BD	Association between mean TRYCAT levels and syndrome-based groups	
Hercules	MDE	n = 255	**Chapter 5:** To study the dynamics of individual depressive symptoms over time and their covariation with other symptoms	Undirected DTW analyses yielding symptom dimensions (i.e. groups of symptoms with similar dynamics over time)	PART 2
		n = 166	**Chapter 6:** To study the temporal direction of changes in depression symptoms	Directed DTW analyses investigating directionality of symptom change over time (i.e. symptoms with high out-strength centrality being more likely to cause fluctuations in other symptoms or symptoms with high in-strength centrality being more susceptible upon fluctuations in other symptoms)	
PROTECT	MDE treated with ECT	n = 65	**Chapter 7:** To examine the evolution of global cognitive function in MDE after treatment with ECT	Group-level changes in cognitive function (LMM) combined with individual-level analyses (RCI)	PART 3
		n = 73	**Chapter 8:** To examine the evolution of a broad array of cognitive tests in MDE after treatment with ECT		

Abbreviations: ECT: electroconvulsive therapy; MDD: major depressive disorder; MDE: major depressive episode; LMM: linear mixed models; TRYCAT: tryptophan catabolites; PROTECT: PRediction Of Treatment response to ElectroConvulsive Therapy study; RCI: reliable change indices

To conclude the thesis, Chapter 9 provides a summary of the main findings and a general discussion.

1.5 Study populations and designs used in this thesis

In this thesis, we have analysed the data from three research projects that assessed patients' mood and cognitive and/or inflammatory variables prospectively at regular time intervals in individuals receiving inpatient treatment for MDE.

1.5.1 The 'HERCULES Routine Outcome Monitoring' study

The main goal of the HERCULES Routine Outcome Monitoring (ROM) study was to investigate clinical characteristics (symptom profiles, treatment outcome and response predictors) and depression symptom dynamics using repeated outcome data obtained in a large cohort of depressed patients who were admitted for treatment to a tertiary psychiatric hospital (University Psychiatric Center Duffel, Belgium). All had an admission diagnosis of a DSM-IV-TR MDE (either within MDD or type-I or type-II BD). To reflect the phenotype of 'real-world' depressed patients as closely as possible, exclusion criteria were restricted to comorbid Mini-International Neuropsychiatric Interview Plus (MINI-Plus) [96] psychotic disorders and alcohol or drugs dependence within 12 months prior to hospitalization. The total sample consisted of 276 patients. The ROM battery was administered at admission and every two weeks throughout the patient's hospital stay (ranging from 2 weeks to 16 months) and included the following instruments: the Hamilton Rating Scale for Depression (HRSD-17) [97], the Beck Depression Inventory (BDI-II) [98], the Manchester Short Assessment of Quality of Life (MANSA) [99] and the Symptom Questionnaire-48 (SQ-48) [100]. Apart from supporting "patient-centered research", ROM also served clinical purposes, with the ROM results serving as a feedback tool for both the clinician and patient.

1.52. The 'Inflammation and Neuroprotection in BD' study

This longitudinal case-control study compares adult participants (age range 18-65 years) with a DSM-IV-TR diagnosis of type-I or type-II BD or a schizoaffective disorder currently experiencing a MDE or (hypo)manic episode and gender- and age-matched healthy controls. Exclusion criteria for both groups included: MINI-Plus version 5.0.0. [96] substance abuse, use of anti-inflammatory drugs within two weeks before screening, acute infection, autoimmune diseases, chronic inflammatory or neurological diseases, pregnancy or breastfeeding, ECT within six months before screening or during follow-up and significant disturbances on the blood screening test at admission. Inpatients were recruited in three psychiatric hospitals in the region of Antwerp. Outpatients were recruited via the patient association "Ups and Downs". The study aimed to investigate the

role of inflammation and tryptophan metabolism in the pathophysiology of BD and BD-related cognitive dysfunction [101–103]. The study cohorts consisted of 35 depressed and 32 (hypo)manic patients with BD and 29 controls. Plasma samples for the quantification of 3-hydroxykynurenine, quinolinic acid and kynurenic acid were drawn at baseline and at four and eight months. Cognitive functioning was assessed at the same three time points using five subtasks the International Society for Bipolar Disorders Battery for assessment of Neurocognition (ISBD-BANC) [104] measuring processing speed (Brief Assessment of Cognition in Schizophrenia-Symbol Coding (Y-BACS), sustained attention (Continuous Performance Task-Identical Pairs (CPT-IP)), verbal memory (Hopkins Verbal Learning Test – Revised (HVLT-R)), working memory (Letter Number Span (LNS)) and response inhibition (Color-Word Interference Test (D-KEFS)).

1.5.3 The 'PROTECT' study

The "Prediction of treatment response to electroconvulsive therapy" (PROTECT) study aimed to investigate the clinical and biological ECT outcome predictors and sought to identify those patients that were at risk of cognitive side effects from ECT [105–108]. Adults with a minimum age of 18 years being treated with ECT for their severe MDE were included. Patients with a co-morbid drug or alcohol dependence or a primary psychotic disorder (as confirmed by the MINI 6.0. [96]) were excluded. Patients were all recruited from the in- or outpatient ward of the University Psychiatric Center Duffel. The study sample consisted of 73 patients of whom 65 completed the full course of ECT. Cognition was assessed prospectively at five time points: prior to ECT (T0), during ECT (before the third session; T1), within one week after ECT completion and again after three (T3) and six (T4) months. The test battery gauged five cognitive domains: global cognition using the Montreal Cognitive Assessment (MoCA)[109], processing speed using the Symbol Digit Substitution Test (SDST)[110], verbal memory using the Hopkins Verbal Learning Test – Revised (HVLT-R) and retrograde amnesia using the Section C of the Kopelman Autobiographic Memory Interview (AMI-C) [111]

References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (5th Ed.)*. Arlington: American Psychiatric Publishing; 2013.
2. McIntyre RS, Cha DS, Soczynska JK, et al. Cognitive deficits and functional outcomes in major depressive disorder: determinants, substrates, and treatment interventions. *Depress Anxiety*. 2013;30(6):515-527. doi:10.1002/DA.22063
3. Rock PL, Roiser JP, Riedel WJ, Blackwell AD. Cognitive impairment in depression: a systematic review and meta-analysis. *Psychol Med*. 2014;44(10):2029-2040. doi:10.1017/S0033291713002535
4. Brissos S, Dias VV, Kapczinski F. Cognitive performance and quality of life in bipolar disorder. *Can J Psychiatry*. 2008;53(8):517-524. doi:10.1177/070674370805300806
5. Ceylan D, Akdede BB, Bora E, et al. Neurocognitive functioning during symptomatic states and remission in bipolar disorder and schizophrenia: A comparative study. *Psychiatry Res*. 2020;292. doi:10.1016/j.psychres.2020.113292
6. James SL, Abate D, Abate KH, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 Diseases and Injuries for 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018;392(10159):1789-1858. doi:10.1016/S0140-6736(18)32279-7
7. Wittchen HU, Jacobi F, Rehm J, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol*. 2011;21(9):655-679. doi:10.1016/j.euroneuro.2011.07.018
8. Wittmund B, Wilms H, Mory C, Angermeyer M. Depressive disorders in spouses of mentally ill patients. *Soc Psychiatry Psychiatr Epidemiol*. 2002;37(4):177-182. doi:10.1007/S001270200012
9. Alonso J, Angermeyer M, Bernert S, et al. 12-Month comorbidity patterns and associated factors in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl*. 2004;109(420):28-37. doi:10.1111/J.1600-0047.2004.00328.X
10. Merikangas KR, Jin R, He J-P, et al. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Arch Gen Psychiatry*. 2011;68(3):241-251. doi:10.1001/archgenpsychiatry.2011.12
11. Vaccarino V, Badimon L, Bremner J, et al. Depression and coronary heart disease: 2018 position paper of the ESC working group on coronary pathophysiology and microcirculation. *Eur Heart J*. 2020;41(17):1687-1696. doi:10.1093/EURHEARTJ/EHY913
12. Benton T, Staab J, Evans D. Medical co-morbidity in depressive disorders. *Ann Clin Psychiatry*. 2007;19(4):289-303. doi:10.1080/10401230701653542
13. Kilbourne A, Cornelius J, Han X, et al. Burden of general medical conditions among individuals with bipolar disorder. *Bipolar Disord*. 2004;6(5):368-373. doi:10.1111/J.1399-5618.2004.00138.X

14. Merikangas K, Jin R, He J, et al. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Arch Gen Psychiatry*. 2011;68(3):241-251. doi:10.1001/ARCHGENPSYCHIATRY.2011.12
15. Plana-Ripoll O, Musliner K, Dalsgaard S, et al. Nature and prevalence of combinations of mental disorders and their association with excess mortality in a population-based cohort study. *World Psychiatry*. 2020;19(3):339-349. doi:10.1002/WPS.20802
16. Greenberg P, Fournier A, Sisitsky T, Pike C, Kessler R. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *J Clin Psychiatry*. 2015;76(2):155-162. doi:10.4088/JCP.14M09298
17. Organisation for Economic Cooperation and Development. (2017). Economic Surveys: Belgium. https://www.oecd.org/eco/surveys/economic-survey-belgium.html. Published 2017. Accessed March 12, 2022.
18. Robins E, Guze S. Establishment of diagnostic validity in psychiatric illness: its application to schizophrenia. *Am J Psychiatry*. 1970;126(7):983-987. doi:10.1176/AJP.126.7.983
19. Joyce PR. The medical model-why psychiatry is a branch of medicine. *Aust N Z J Psychiatry*. 1980;14(4):269-278.
20. *World Health Organization. (2018). International Classification of Diseases for Mortality and Morbidity Statistics (11th Revision)*.
21. Smoller JW, Kendler K, Craddock N, et al. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet (London, England)*. 2013;381(9875):1371-1379. doi:10.1016/S0140-6736(12)62129-1
22. Hek K, Demirkan A, Lahti J, et al. A genome-wide association study of depressive symptoms. *Biol Psychiatry*. 2013;73(7):667-678. doi:10.1016/J.BIOPSYCH.2012.09.033
23. Lewis CM, Ng MY, Butler AW, Cohen-Woods S, Uher R, Pirlo K, Weale ME, Schosser A, Paredes UM, Rivera M, Craddock N, Owen MJ, Jones L, Jones I, Korszun A, Aitchison KJ, Shi J, Quinn JP, Mackenzie A, Vollenweider P, Waeber G, Heath S, Lathrop M, Muglia P, B MP. Genome-wide association study of major recurrent depression in the U.K. population. *Am J Psychiatry*. 2010;167(8):949-957. doi:10.1176/APPI.AJP.2010.09091380
24. J S, JB P, JA K, et al. Genome-wide association study of recurrent early-onset major depressive disorder. *Mol Psychiatry*. 2011;16(2):193-201. doi:10.1038/MP.2009.124
25. Kapur S, Phillips A, Insel T. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry*. 2012;17(12):1174-1179. doi:10.1038/MP.2012.105
26. van Eeden WA, Luo C, van Hemert AM, et al. Predicting the 9-year course of mood and anxiety disorders with automated machine learning: A comparison between auto-sklearn, naïve Bayes classifier, and traditional logistic regression. *Psychiatry Res*. 2021;299:113823. doi:10.1016/J.PSYCHRES.2021.113823
27. Nelson B, McGorry PD, Wichers M, Wigman JTW, Hartmann JA. Moving from static to dynamic models of the onset of mental disorder a review. *JAMA Psychiatry*. 2017;74(5):528-534. doi:10.1001/jamapsychiatry.2017.0001

28. Ghaemi SN. After the failure of DSM: clinical research on psychiatric diagnosis. *World Psychiatry*. 2018;17(3):301. doi:10.1002/WPS.20563
29. Bringmann LF, Eronen MI. Don't Blame the model: Reconsidering the network approach to psychopathology. *Psychol Rev*. 2018;125(4):606-615. doi:10.1037/rev0000108
30. Fried EI. Problematic assumptions have slowed down depression research: Why symptoms, not syndromes are the way forward. *Front Psychol*. 2015;6(MAR). doi:10.3389/fpsyg.2015.00309
31. Borsboom D. Psychometric perspectives on diagnostic systems. *J Clin Psychol*. 2008;64(9):1089-1108. doi:10.1002/jclp.20503
32. Borsboom D, Mellenbergh GJ, Van Heerden J. The Theoretical Status of Latent Variables. *Psychol Rev*. 2003;110(2):203-219. doi:10.1037/0033-295X.110.2.203
33. Schmittmann VD, Cramer AOJ, Waldorp LJ, Epskamp S, Kievit RA, Borsboom D. Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas Psychol*. 2013;31(1):43-53. doi:10.1016/j.newideapsych.2011.02.007
34. Lynch SJ, Sunderland M, Newton NC, Chapman C. A systematic review of transdiagnostic risk and protective factors for general and specific psychopathology in young people. *Clin Psychol Rev*. 2021;87:102036. doi:10.1016/J.CPR.2021.102036
35. Feola B, Armstrong K, Flook E, Woodward N, Heckers S, Blackford J. Evidence for inhibited temperament as a transdiagnostic factor across mood and psychotic disorders. *J Affect Disord*. 2020;274:995-1003. doi:10.1016/J.JAD.2020.05.119
36. Navarro V, Bal Gastó C, Torres X, et al. Continuation/Maintenance Treatment with Nortriptyline Versus Combined Nortriptyline and ECT in Late-Life Psychotic Depression: A Two-Year Randomized Study. doi:10.1097/JGP.0b013e318170a6fa
37. EI F, RM N. Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. *J Affect Disord*. 2015;172:96-102. doi:10.1016/J.JAD.2014.10.010
38. Fava M, Rush AJ, Trivedi MH, et al. Background and rationale for the sequenced treatment alternatives to relieve depression (STAR*D) study. *Psychiatr Clin North Am*. 2003;26(2):457-494. doi:10.1016/S0193-953X(02)00107-7
39. Rush AJ, Fava M, Wisniewski SR, et al. Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design. *Control Clin Trials*. 2004;25(1):119-142. doi:10.1016/S0197-2456(03)00112-0
40. RC K, KR M, PS W. Prevalence, comorbidity, and service utilization for mood disorders in the United States at the beginning of the twenty-first century. *Annu Rev Clin Psychol*. 2007;3:137-158. doi:10.1146/ANNUREV.CLINPSY.3.022806.091444
41. Kupfer DJ, First MB, Regier DA. *A Research Agenda for DSM-V*. (Association AP, ed.).; 2002. www.psych.org. Accessed October 22, 2021.
42. Kessler RC, Chiu WT, Demler O, Walters EE. Prevalence, Severity, and Comorbidity of 12-Month DSM-IV Disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):617-627. doi:10.1001/ARCHPSYC.62.6.617

43. Lamers F, van Oppen P, Comijs H, et al. Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). *J Clin Psychiatry*. 2011;72(3):342-348. doi:10.4088/JCP.10M06176BLU
44. Lai H, Cleary M, Sitharthan T, Hunt G. Prevalence of comorbid substance use, anxiety and mood disorders in epidemiological surveys, 1990-2014: A systematic review and meta-analysis. *Drug Alcohol Depend*. 2015;154:1-13. doi:10.1016/J.DRUGALCDEP.2015.05.031
45. N H, E H, P K. Categories versus dimensions in personality and psychopathology: a quantitative review of taxometric research. *Psychol Med*. 2012;42(5):903-920. doi:10.1017/S0033291711001966
46. RF K, TM P. Toward a dimensional and psychometrically-informed approach to conceptualizing psychopathology. *Behav Res Ther*. 2002;40(5):485-499. doi:10.1016/S0005-7967(02)00016-5
47. Borsboom D, Cramer AOJ. Network Analysis: An Integrative Approach to the Structure of Psychopathology. *Annu Rev Clin Psychol*. 2013;9(1):91-121. doi:10.1146/annurev-clinpsy-050212-185608
48. Fried EI, Nesse RM. Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. *BMC Med*. 2015;13(1):72. doi:10.1186/s12916-015-0325-4
49. Fried EI, Nesse RM, Zivin K, Guille C, Sen S. Depression is more than the sum score of its parts: individual DSM symptoms have different risk factors. *Psychol Med*. 2014;44(10):2067-2076. doi:10.1017/S0033291713002900
50. Lux V, Kendler KS. Deconstructing major depression: a validation study of the DSM-IV symptomatic criteria. *Psychol Med*. 2010;40(10):1679. doi:10.1017/S0033291709992157
51. Fried EI, Nesse RM. The impact of individual depressive symptoms on impairment of psychosocial functioning. *PLoS One*. 2014;9(2). doi:10.1371/journal.pone.0090311
52. D R, F L, J S, R de G, AT B, BW P. Course trajectories of unipolar depressive disorders identified by latent class growth analysis. *Psychol Med*. 2012;42(7):1383-1396. doi:10.1017/S0033291711002509
53. Chekroud AM, Gueorguieva R, Krumholz HM, Trivedi MH, Krystal JH, Mccarthy G. Reevaluating the Efficacy and Predictability of Antidepressant Treatments A Symptom Clustering Approach Supplemental content. *JAMA Psychiatry*. 2017;74(4):370-378. doi:10.1001/jamapsychiatry.2017.0025
54. Hieronymus F, Emilsson JF, Nilsson S, Eriksson E. Consistent superiority of selective serotonin reuptake inhibitors over placebo in reducing depressed mood in patients with major depression. *Mol Psychiatry*. 2016;21(4):523-530. doi:10.1038/mp.2015.53
55. Briganti G, Linkowski P. Exploring network structure and central items of the Narcissistic Personality Inventory. *Int J Methods Psychiatr Res*. 2020;29(1). doi:10.1002/MPR.1810
56. Insel T, Cuthbert B, Garvey M, et al. Research Domain Criteria (RDoC): Toward a new classification framework for research on mental disorders. *Am J Psychiatry*. 2010;167(7):748-751. doi:10.1176/appi.ajp.2010.09091379

57. Wichers M, Wigman JTW, Bringmann LF, de Jonge P. Mental disorders as networks: some cautionary reflections on a promising approach. *Soc Psychiatry Psychiatr Epidemiol*. 2017;52(2):143-145. doi:10.1007/s00127-016-1335-z
58. Borsboom D, Cramer AOJ, Schmittmann VD, Epskamp S, Waldorp LJ. The Small World of Psychopathology. *PLoS One*. 2011;6(11):e27407. doi:10.1371/JOURNAL.PONE.0027407
59. Wichers M. The dynamic nature of depression: A new micro-level perspective of mental disorder that meets current challenges. *Psychol Med*. 2014;44(7):1349-1360. doi:10.1017/S0033291713001979
60. Olbert CM, Gala GJ, Tupler LA. Quantifying heterogeneity attributable to polythetic diagnostic criteria: theoretical framework and empirical application. *J Abnorm Psychol*. 2014;123(2):452-462. doi:10.1037/A0036068
61. Molenaar PCM. A Manifesto on Psychology as Idiographic Science: Bringing the Person Back Into Scientific Psychology, This Time Forever. *http://dx.doi.org/101207/s15366359mea0204_1*. 2009;2(4):201-218. doi:10.1207/S15366359MEA0204_1
62. Beltz AM, Wright AGC, Sprague BN, Molenaar PCM. Bridging the Nomothetic and Idiographic Approaches to the Analysis of Clinical Data. *Assessment*. 2016;23(4):447-458. doi:10.1177/1073191116648209
63. MJ K, BN C. The NIMH Research Domain Criteria Initiative: Background, Issues, and Pragmatics. *Psychophysiology*. 2016;53(3):286-297. doi:10.1111/PSYP.12518
64. Kendler KS, Zachar P, Craver C. What kinds of things are psychiatric disorders? *Psychol Med*. 2011;41(6):1143-1150. doi:10.1017/S0033291710001844
65. Borsboom D. A network theory of mental disorders. *World Psychiatry*. 2017;16(1):5-13. doi:10.1002/wps.20375
66. Cramer AOJ, Van Borkulo CD, Giltay EJ, et al. Major depression as a complex dynamic system. *PLoS One*. 2016;11(12). doi:10.1371/journal.pone.0167490
67. Mandell AJ, Selz KA. Dynamical systems in psychiatry: now what? *Biol Psychiatry*. 1992;32(4):299-301. doi:10.1016/0006-3223(92)90034-W
68. Globus GG, Arpaia JP. Psychiatry and the new dynamics. *Biol Psychiatry*. 1994;35(5):352-364. doi:10.1016/0006-3223(94)90039-6
69. Odgers CL, Mulvey EP, Skeem JL, Gardner W, Lidz CW, Schubert C. Capturing the ebb and flow of psychiatric symptoms with dynamical systems models. *Am J Psychiatry*. 2009;166(5):575-582. doi:10.1176/APPI.AJP.2008.08091398/ASSET/IMAGES/LARGE/U412T1.JPEG
70. Robinaugh DJ, Hoekstra RHA, Toner ER, Borsboom D. The network approach to psychopathology: A review of the literature 2008-2018 and an agenda for future research. *Psychol Med*. 2020;50(3):353-366. doi:10.1017/S0033291719003404
71. Bringmann LF, Ferrer E, Hamaker EL, Borsboom D, Tuerlinckx F. Modeling Nonstationary Emotion Dynamics in Dyads using a Time-Varying Vector-Autoregressive Model. *Multivariate Behav Res*. 2018;53(3):293-314. doi:10.1080/00273171.2018.1439722

72. Fisher AJ, Reeves JW, Lawyer G, Medaglia JD, Rubel JA. Exploring the idiographic dynamics of mood and anxiety via network analysis. *J Abnorm Psychol*. 2017;126(8):1044-1056. doi:10.1037/abn0000311
73. KM D, P G, LJ R, et al. Prevalence of cognitive impairment in major depression and bipolar disorder. *Bipolar Disord*. 2018;20(3):260-274. doi:10.1111/BDI.12602
74. HL K, FF E, M V, LV K, K M. Neurocognitive heterogeneity in patients with bipolar disorder and their unaffected relatives: associations with emotional cognition. *Psychol Med*. 2021;51(4):668-679. doi:10.1017/S0033291719003738
75. Obbels J, Verwijk E, Vansteelandt K, et al. Long-term neurocognitive functioning after electroconvulsive therapy in patients with late-life depression. *Acta Psychiatr Scand*. 2018;138(3):223-231. doi:10.1111/acps.12942
76. Dybedal GS, Tanum L, Sundet K, Gaarden TL, Bjølseth TM. Cognitive side-effects of electroconvulsive therapy in elderly depressed patients. *Clin Neuropsychol*. 2014;28(7):1071-1090. doi:10.1080/13854046.2014.958536
77. Nuninga JO, Claessens TFI, Somers M, et al. Immediate and long-term effects of bilateral electroconvulsive therapy on cognitive functioning in patients with a depressive disorder. *J Affect Disord*. 2018;238:659-665. doi:10.1016/j.jad.2018.06.040
78. Yatham LN, Kennedy SH, Parikh S V, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) 2018 guidelines for the management of patients with bipolar disorder. *Bipolar Disord*. 2018;20(2):97-170. doi:10.1111/bdi.12609
79. Goodwin GM, Haddad PM, Ferrier IN, et al. Evidence-based guidelines for treating bipolar disorder: Revised third edition recommendations from the British Association for Psychopharmacology. *J Psychopharmacol*. 2016;30(6):495-553. doi:10.1177/0269881116636545
80. Mitchell PB, Wilhelm K, Parker G, Austin MP, Rutgers P, Malhi GS. The clinical features of bipolar depression: a comparison with matched major depressive disorder patients. *J Clin Psychiatry*. 2001;62(3):212-216; quiz 217.
81. Miller AH, Maletic V, Raison CL. Inflammation and Its Discontents: The Role of Cytokines in the Pathophysiology of Major Depression. *Biol Psychiatry*. 2009;65(9):732-741. doi:10.1016/J.BIOPSYCH.2008.11.029
82. Van Den Ameele S, Van Diermen L, Staels W, et al. The effect of mood-stabilizing drugs on cytokine levels in bipolar disorder: A systematic review. *J Affect Disord*. 2016;203:364-373. doi:10.1016/j.jad.2016.06.016
83. Wang YL, Han QQ, Gong WQ, et al. Microglial activation mediates chronic mild stress-induced depressive- and anxiety-like behavior in adult rats. *J Neuroinflammation*. 2018;15(1). doi:10.1186/S12974-018-1054-3
84. Haarman BCMB, Riemersma-Van der Lek RF, de Groot JC, et al. Neuroinflammation in bipolar disorder – A [11C]-(R)-PK11195 positron emission tomography study. *Brain Behav Immun*. 2014;40:219-225. doi:10.1016/J.BBI.2014.03.016

85. Dantzer R, O'Connor JC, Lawson MA, Kelley KW. Inflammation-associated depression: from serotonin to kynurenine. *Psychoneuroendocrinology*. 2011;36(3):426-436. doi:10.1016/J.PSYNEUEN.2010.09.012
86. Savitz J. The kynurenine pathway: a finger in every pie. *Mol Psychiatry*. 2020;25(1):131-147. doi:10.1038/s41380-019-0414-4
87. Schwarcz R, Bruno JP, Muchowski PJ, Wu HQ. Kynurenines in the mammalian brain: When physiology meets pathology. *Nat Rev Neurosci*. 2012;13(7):465-477. doi:10.1038/nrn3257
88. Rosa-Neto P, Diksic M, Okazawa H, et al. Measurement of brain regional alpha-[11C]methyl-L-tryptophan trapping as a measure of serotonin synthesis in medication-free patients with major depression. *Arch Gen Psychiatry*. 2004;61(6):556-563. doi:10.1001/ARCHPSYC.61.6.556
89. Vaswani M, Linda FK, Ramesh S. Role of selective serotonin reuptake inhibitors in psychiatric disorders: a comprehensive review. *Prog Neuropsychopharmacol Biol Psychiatry*. 2003;27(1):85-102. doi:10.1016/S0278-5846(02)00338-X
90. Arango V, Underwood MD, Mann JJ. Chapter 35 Serotonin brain circuits involved in major depression and suicide. *Prog Brain Res*. 2002;136:443-453. doi:10.1016/S0079-6123(02)36037-0
91. Fujigaki H, Yamamoto Y, Saito K. L-Tryptophan-kynurenine pathway enzymes are therapeutic target for neuropsychiatric diseases: Focus on cell type differences. *Neuropharmacology*. 2017;112(Pt B):264-274. doi:10.1016/J.NEUROPHARM.2016.01.011
92. Guillemin GJ. Quinolinic acid, the inescapable neurotoxin. *FEBS J*. 2012;279(8):1356-1365. doi:10.1111/J.1742-4658.2012.08485.X
93. Pu J, Liu Y, Zhang H, et al. An integrated meta-analysis of peripheral blood metabolites and biological functions in major depressive disorder. *Mol Psychiatry*. 2021;26(8):4265-4276. doi:10.1038/s41380-020-0645-4
94. Marx W, McGuinness AJ, Rocks T, et al. The kynurenine pathway in major depressive disorder, bipolar disorder, and schizophrenia: a meta-analysis of 101 studies. *Mol Psychiatry*. 2020. doi:10.1038/s41380-020-00951-9
95. Ogyu K, Kubo K, Noda Y, et al. Kynurenine pathway in depression: A systematic review and meta-analysis. *Neurosci Biobehav Rev*. 2018;90(February):16-25. doi:10.1016/j.neubiorev.2018.03.023
96. Sheehan D V, Lecrubier Y, Sheehan KH, et al. 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. *J Clin Psychiatry*. 1998;59 Suppl 2:22-33;quiz 34-57. http://www.ncbi.nlm.nih.gov/pubmed/9881538. Accessed November 11, 2018.
97. HAMILTON M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23:56-62. doi:10.1136/jnnp.23.1.56
98. Beck AT, Steer RA BO. *Beck Depression Inventory Manual, 2nd Ed*. Psychological Corporation, San Antonio, TX; 1996.
99. Priebe S, Huxley P, Knight S, Evans S. Application and Results of the Manchester Short Assessment of Quality of Life (Mansa). *Int J Soc Psychiatry*. 1999;45(1):7-12. doi:10.1177/002076409904500102

100. Carlier I, Schulte-Van Maaren Y, Wardenaar K, et al. Development and validation of the 48-item Symptom Questionnaire (SQ-48) in patients with depressive, anxiety and somatoform disorders. *Psychiatry Res*. 2012;200(2-3):904-910. doi:10.1016/j.psychres.2012.07.035
101. van den Ameele S, Coppens V, Schuermans J, et al. Neurotrophic and inflammatory markers in bipolar disorder: A prospective study. *Psychoneuroendocrinology*. 2017;84(May):143-150. doi:10.1016/j.psyneuen.2017.07.003
102. van den Ameele S, Fuchs D, Coppens V, et al. Markers of Inflammation and Monoamine Metabolism Indicate Accelerated Aging in Bipolar Disorder. *Front Psychiatry*. 2018;9:250. doi:10.3389/fpsyt.2018.00250
103. van den Ameele S, van Nuijs AL, Lai FY, et al. A mood state-specific interaction between kynurenine metabolism and inflammation is present in bipolar disorder. *Bipolar Disord*. 2020;22(1):59-69. doi:10.1111/bdi.12814
104. Yatham LN, Torres IJ, Malhi GS, et al. The International Society for Bipolar Disorders–Battery for Assessment of Neurocognition (ISBD-BANC). *Bipolar Disord*. 2010;12(4):351-363. doi:10.1111/J.1399-5618.2010.00830.X
105. van Diermen L, Vanmarcke S, Walther S, et al. Can psychomotor disturbance predict ect outcome in depression? *J Psychiatr Res*. 2019;117(May):122-128. doi:10.1016/j.jpsychires.2019.07.009
106. van Diermen L, Hebbrecht K, Schrijvers D, Sabbe BCG, Fransen E, Birkenhäger TK. The Maudsley Staging Method as predictor of electroconvulsive therapy effectiveness in depression. *Acta Psychiatr Scand*. 2018;138(6). doi:10.1111/acps.12962
107. van Diermen L, Walther S, Cools O, et al. Observer-rated retardation but not agitation corresponds to objective motor measures in depression. *Acta Neuropsychiatr*. 2018;30(06):359-364. doi:10.1017/neu.2018.21
108. L van D, E P, R V der M, et al. Toward Targeted ECT: The Interdependence of Predictors of Treatment Response in Depression Further Explained. *J Clin Psychiatry*. 2020;82(1). doi:10.4088/JCP.20M13287
109. Nasreddine ZS, Phillips NA, BÃ©dirian V, et al. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *J Am Geriatr Soc*. 2005;53(4):695-699. doi:10.1111/j.1532-5415.2005.53221.x
110. Jaeger J. Digit symbol substitution test. *J Clin Psychopharmacol*. 2018;38(5):513-519. doi:10.1097/JCP.0000000000000941
111. Kopelman MD, Wilson BA, Baddeley AD. The autobiographical memory interview: a new assessment of autobiographical and personal semantic memory in amnesic patients. *J Clin Exp Neuropsychol Off J Int Neuropsychol Soc*. 1989;11(5):724-744. doi:10.1080/01688638908400928

CHAPTER 2

Symptom profile and clinical course of inpatients with unipolar versus bipolar depression

Hebbrecht K, Stuivenga M, Birkenhäger T, van der Mast RC, Sabbe B, Giltay EJ.

Neuropsychobiology. 2020;79(4-5):313-323.

Abstract

Background: Although differences in symptom profiles and outcome between depressive patients with an underlying major depressive disorder (MDD) and bipolar depression (BD) have been reported, studies with sequential short-interval assessments in a real-life inpatient setting are scarce.

Objectives: To examine potential differences in symptom profile and course of depressive symptomatology in depressive inpatients with underlying MDD and BD.

Methods: A cohort of 276 consecutive inpatients with MDD (n = 224) or BD (n = 52) was followed during their hospitalization using routine outcome monitoring (ROM), which included a structured diagnostic interview at baseline (Mini-International Neuropsychiatric Interview Plus (MINIPlus) and repeated 17-item Hamilton Depression Rating Scale every 2 weeks. MDD and BD were compared regarding their symptom profiles and time to response and remission. Furthermore, the concordance between the MINI-Plus and clinical diagnosis was analyzed.

Results: Patients were on average 52 and 47 years old in the MDD and BD group, respectively, and 66 versus 64% were female. Compared to patients with BD, patients with MDD scored higher on weight loss ($p = 0.02$), whereas the BD group showed a higher long- term likelihood of response (hazard ratio = 1.93, 95% confidence interval 1.16–3.20, p for interaction with time = 0.04). Although the same association was seen for remission, the interaction with time was not significant ($p = 0.48$). Efficiency between the MINI-Plus and clinical diagnosis of BD was high (0.90), suggesting that the MINI-Plus is an adequate ROM diagnostic tool.

Conclusions: In routine clinical inpatient care, minor differences in the symptom profile and the course of depressive symptomatology may be helpful in distinguishing MDD and BD, particularly when using sequential ROM assessments.

Introduction

Patients who are admitted to a psychiatric hospital due to depressive symptoms often suffer from an underlying major depressive disorder (MDD) or bipolar depression (BD). Since depressive episodes are often the initial manifestation of bipolar disorder and multiple depressive episodes can precede the first episode of (hypo)mania, a bipolar disorder is often misidentified as a unipolar MDD [1-3].

Previous studies have not found consistent characteristic signs distinguishing unipolar from bipolar major depression, although some clinical features may discriminate between MDD and BD [4,5]. BD, relative to MDD, has been more frequently characterized by atypical features, such as hypersomnia or hyperphagia [6-8]; melancholic symptoms, such as anhedonia, diurnal variation, and unvarying mood [6,9]; and psychotic features [6,10–12]. Furthermore, psychomotor retardation is especially more frequent in patients with bipolar I disorder experiencing a depressive episode than in patients with unipolar depression [6,9,13]. Meanwhile, MDD often coincides with weight loss [14], initial insomnia [6,15], and more anxiety and agitation than BD [6,12,16,17].

Differences in the clinical course between MDD and BD have also been found. Two retrospective comparative studies, with an inclusion rate of 108/50 MDD/BD in- and outpatients [18] and 59/96 MDD/BD inpatients [19], have shown that BD was associated with a faster onset of depressive symptoms than MDD. We also found six prospective observational studies that investigated the course of bipolar and unipolar mood disorders in a naturalistic setting [20–25], but all used different operational criteria in terms of diagnostics and outcome. Furthermore, all but one study [20] focused on the course of bipolar disorder as a whole (i.e., comparing mania and depression in bipolar disorder with depression in unipolar depressive disorder), compared to focusing only on the course of the depressive episode. These studies quite consistently reported a longer duration of the depressive episode in MDD compared to BD [20,22,23]; however, one study [24] contradicted these results. The study by Furukuwa et al. [20] specifically compared improvement rates of depressive symptoms between MDD and BD patients under naturalistic conditions but failed to find a significant difference in median time to recovery between MDD (n = 95) and BD (n = 6). Our study is unique in terms of the short intervals (every 2 weeks) between assessments during admission that were routinely incorporated into clinical practice (routine outcome monitoring; ROM).

When systematically comparing MDD and BD patients, ROM could be a helpful tool to diagnose and to clinically assess depression severity [26]. ROM entails the collection of clinical data at baseline and at regular time intervals thereafter in order to monitor disease severity as well as clinical improvement during treatment. ROM may provide feedback to both

the clinician and the patient and effectuate "patient-centered research" [27,28]. Standardized diagnostic interviews (SDIs), such as the Composite International Diagnostic Interview (CIDI) [29] or the shorter Mini-International Neuropsychiatric Interview Plus (MINI-Plus) [30], are often a part of ROM, with which Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) diagnoses can be established. However, the agreement between standardized diagnoses (often used as a diagnostic method in research) and clinical diagnoses has not been extensively studied in inpatient samples [31], which restricts the generalizability of research results to clinical practice. In a large outpatient study by Verhoeven et al. [32], the concordance indices between the MINI-Plus and clinical diagnoses for diverse mood and anxiety disorders were only of moderate strength.

Our study had several aims. First, we compared baseline clinical features of inpatients diagnosed with MDD versus BD. Second, we evaluated the diagnostic concordance between the MINI and the clinical diagnoses among both groups. Third, we compared the differences in the time to response and remission during the admission period between both groups. We hypothesized that BD, compared to MDD, would be associated with a specific symptom profile, with good concordance indices between the MINI-Plus and clinical diagnoses and with higher and faster response rates.

Methods

Setting and population

Patients consecutively admitted to a tertiary psychiatric hospital in Duffel, Belgium and fulfilling the MINI-Plus diagnosis of MDD or BD were included in the study. In order to obtain a representative sample of participants, exclusion criteria were minimal. We excluded patients with comorbid MINI-Plus psychotic disorders (including schizoaffective disorder) and dependence on alcohol or drugs within 12 months prior to hospitalization, as well as patients with insufficient mastery of the Dutch language.

Routine outcome monitoring

ROM has been an integral part of clinical practice at the Psychiatric Hospital of Duffel since 2015, especially with respect to the management of depression. All admitted patients fulfilling the MINI diagnosis of MDD or BD were routinely assessed with a variety of psychometric tests at admission and every 2 weeks during hospitalization. ROM was carried out by a trained study psychologist or psychiatrist, both of whom were not involved in the clinical care of that patient.

ROM measurements were done at baseline and approximately every 2 weeks thereafter during the clinical admission. The first ROM assessment (preferably took place on the

first day of admission) consisted of a standardized diagnostic interview (Dutch version of the MINI-Plus) [30,33], an observer-rated scale (Hamilton Depression Rating Scale, HDRS-17) [34,35], and the self-report Beck Depression Inventory-II (BDI-II) questionnaire [36]. The first assessment was completed in a face-to-face interview, whereas the latter two were filled out by the patient on a computer (VitalHealth Software, QuestManager). Sociodemographic and clinical characteristics were gathered at baseline. During the subsequent ROM assessments, every 2 weeks, the HDRS-17 and the BDI-II were administered.

The HDRS-17 consists of 17 items on a Likert scale, ranging from either 0 to 4 or 0 to 2. The internal consistency is high and, considering its frequent clinical use, it has a good test–retest reliability [35]. In order to improve interrater reliability, Hamilton Depression Rating Scale training sessions were organized every three months, during which videorecorded interviews with patients were rated and discussed to reach consensus. The BDI-II consists of 21 multiple-choice, self-report items, which are all related to symptoms of depression. There are four statements per item, ranging from 0 (symptom not present) to 3 (symptom very severe). The total score ranges from 0 to 63 [36]. The BDI-II has a good internal consistency (around 0.9) and a high convergent validity with the HDRS-17 [37].

Treatment

This being an observational study, inpatients received treatment with antidepressants and psychotherapy as usual in our hospital. The treating psychiatrists determined all treatment decisions based on clinical judgment—independent of our study. Treatments during hospitalization were based on evidence-based guidelines and consisted of pharmacotherapy, (group)psychotherapy, or a combination of both. These guidelines for diagnosis and treatment were formulated by the Dutch Association of Psychiatry, often in association with the associations of psychology and general practitioners (www.trimbos.nl,www.nvvp.net). Medications were grouped into five categories: antidepressants, mood stabilizers, antipsychotics, benzodiazepines, and stimulants.

Diagnosis of mood disorder

We used the MINI-Plus (5.0.0) at hospital admission to identify all possible patients with DSM-IV MDD or BD. All MINI interviewers were psychiatrists or psychologists who received a half-day MINI training session. The MINI-Plus is a semistructured diagnostic psychiatric interview for DSM-IV and ICD-10 psychiatric disorders [30,33]; it has good psychometric properties, with interrater reliability of 1.00 and test–retest reliability of 0.87 for MDD, and 0.89 and 0.63 for lifetime mania, respectively [38]. It has adequate validity compared to the CIDI and the Structured Clinical Interview for DSM-III (SCID-patient rated) [39,40]. We used the validated Dutch version of the MINI Plus [33].

Clinical diagnoses were established by treating psychiatrists or supervised residents-in-training working on the hospital ward. Clinicians used any source of possibly relevant information in order to provide a clinical diagnosis, for example, unstructured/open diagnostic interviews or family interviews. For insurance purposes, it is mandatory that clinical diagnoses, described in DSM-IV criteria, are available in the electronic medical record (EMR) of every hospitalized patient. In order to verify this clinical diagnosis, two authors (K.H. and M.S) independently compared the DSM-IV clinical diagnosis to other available diagnostic information in the EMRs. In case of inconsistencies between the DSM-IV diagnosis and diagnostic information in the EMR, the treating psychiatrist was consulted.

Time to response and remission

The main clinical outcomes assessing the course of the depressive episode were time to response and remission of depressive symptoms during admission. We defined response as a 50% reduction (or more) of the HDRS-17 at the first follow-up assessment compared to the baseline assessment. We defined remission as scoring seven or less on the HDRS-17 at the first follow-up assessment. When a patient failed to reach criteria for response and remission during his or her hospitalization period, that patient was censored at the last assessment (at discharge).

Statistical analysis

We compared baseline demographic and illness characteristics between the MDD and BD group using chi-squared tests for categorical data and an analysis of variance (ANOVA) for continuous measures. Whe investigated the differences in depressive symptom profile between MDD and BD, adjustments were made for age, sex, level of education and use of psychotropic drugs (5 classes of psychotropic drugs) using multiple regression analysis. We evaluated the agreement between the MINI and clinical diagnosis using concordance statistics such as sensitivity, specificity, positive and negative predictive values (PPV and NPV), the area below the receiver operator characteristic curve (AUC), and efficiency. We used the clinical diagnosis as the "gold standard," being a realistic representation of daily clinical practice, and investigated the diagnostic power of the MINI correctly diagnosing bipolar disorder at baseline (i.e., at admission to the psychiatric hospital). In this way, sensitivity is the ability of the MINI-Plus to correctly diagnose a patient (with reference to clinical diagnosis), whereas specificity is the ability of the MINI to correctly diagnose a patient not having a specific clinical diagnosis. PPV refers to the probability that a diagnosis of bipolar disorder on the MINI correctly corresponds to an individual diagnosed with bipolar disorder. NPV is the probability that the absence of a diagnosis of bipolar disorder on the MINI actually corresponds to a person not having the clinical diagnosis. Efficiency is the number of cases correctly identified divided by the sample size/relative number of correctly identified cases.

Time-to-response and time-to-remission data were analyzed using the Kaplan–Meier (KM) method and log-rank tests. The time of the first ROM assessment was the origin of the survival time (i.e., t = 0). Hazard ratios (HR) and 95% confidence intervals (CI) were computed using Cox proportional hazard models. The proportional hazard assumption was tested using the loglog graphical method, for which continuous variables were provisionally categorized. Multivariable Cox models included age, sex, level of education, number of episodes, duration of illness and whether the current episode was the index episode or not. We made no adjustments for multiple comparisons because we only tested a priori formulated hypotheses. We also added a time-dependent covariate ($\log_e$ time) to the Cox proportional hazard model in case of an indication of a violation of the proportional hazard assumption [41]. Two-sided p values < .05 were considered statistically significant. Data were analyzed using IBM SPSS Version 25.0 (IBM Corp.).

Results

Baseline sample characteristics

The study sample consisted of 276 patients who met the MINI-Plus diagnostic criteria of MDD (*n* = 224) or BD (*n* = 52). As shown in Table 1, there were no significant differences in sociodemographic characteristics between the MDD and BD groups-except for a difference in age (higher in MDD vs. BD), and more patients were married in the MDD group.

Table 1. Characteristics of 276 consecutive inpatients with unipolar (MDD) and bipolar depression (BD)

	MINI-Plus MDD (*n* = 224)	MINI-Plus BD (*n* = 52)	*p* value[d]
Female (%)	147 (65.6)	33 (63.5)	0.77
Age in years, mean (*SD*)	52.0 ± 15.6	46.8 ± 13.9	0.03
Education[a,b] (n, %)			0.85[e]
- Lower	36 (16.1)	8 (15.7)	
- Intermediate	107 (48.0)	26 (51.0)	
- Higher	80 (35.9)	17 (33.3)	
Work status			0.12
- Unemployed	131 (58.5)	34 (65.4)	
- Employed	83 (37.1)	13 (25.0)	
- Other (student/voluntary service)	10 (4.5)	5 (9.6)	
Marital status			0.01
- Married	103 (46.0)	12 (23.1)	
- Divorced/widowed	55 (24.6)	18 (34.6)	
- Never married	66 (29.5)	22 (42.3)	
Living situation[c]			0.35
- Living alone	65 (29.0)	19 (36.5)	
- Living with partner	88 (39.3)	15 (28.8)	
- Living with family	71 (31.7)	19 (34.6)	
Duration of illness, years[f]	10.2 ± 11.2 (0.08-48)	15.5 ± 11.1 (0.25-40)	0.003
Index (first) depressive episode (n, %)	60 (26.8)	3 (5.8)	0.001
History of 4 or more depressive episodes (n, %)	54 (24.1)	28 (53.8)	< 0.001
Lifetime substance abuse/dependency[g]			
- Alcohol	20 (8.9)	8 (15.4)	0.16
- Drugs (THC, hard drugs, benzodiazepines)	4 (1.8)	1 (1.9)	0.95
Suicidality[g]	103 (46.0)	22 (42.3)	0.63
Melancholic features	145 (64.7)	25 (48.1)	0.03
ROM baseline total scores:			
- HDRS-17	21.0 ± 4.7	19.9 ± 4.0	0.11
- BDI-II	33.9 ± 9.2	32.9 ± 10.1	0.50

Values are presented as n (%) or mean ± standard deviation with or without range in parentheses. BD, bipolar depression; BDI-II, Beck Depression Inventory-II; HDRS-17, 17-item Hamilton Depression Rating Scale; MDD, major depressive disorder; MINI-Plus, Mini-International Neuropsychiatric Interview Plus; ROM, routine outcome monitoring. [a] p values of t test, $\chi 2$ test, or one-way analysis of variance; p values <0.05 are statistically significant. [b] Lower education: general basic education only; intermediate education: middle vocational education; higher education: higher vocational education or university. [c] Two missing values for education. [d] Linear- by-linear association. [e] Living alone includes living in a home for elderly or a convent. [f] Fifteen missing values for duration of illness. [g] Investigated using the MINI-Plus modules on substance abuse and suicidality.

Although there was no difference in the overall severity of depression according to the HDRS-17 total score, patients with MDD were more likely to manifest hypochondriasis and weight loss (HDRS-17 Subitems 15 and 16) (Fig. 1a, unadjusted model). When we adjusted these analyses for age, sex, level of education, antidepressants, mood stabilizers, antipsychotics, lithium, and quetiapine using multiple regression analysis, only weight loss remained statistically significant (p = 0.02) (Fig.1b, adjusted model). The adjustment declined from 0.340 to 0.268 (i.e., 29% decline) for weight loss. The presence of a MINI-Plus depression with melancholic features was more common in the MDD group than in the BD group (Table 1). The presence of a MINI-Plus depression with melancholic features was more common in the MDD group than in the BD group (Table 1).

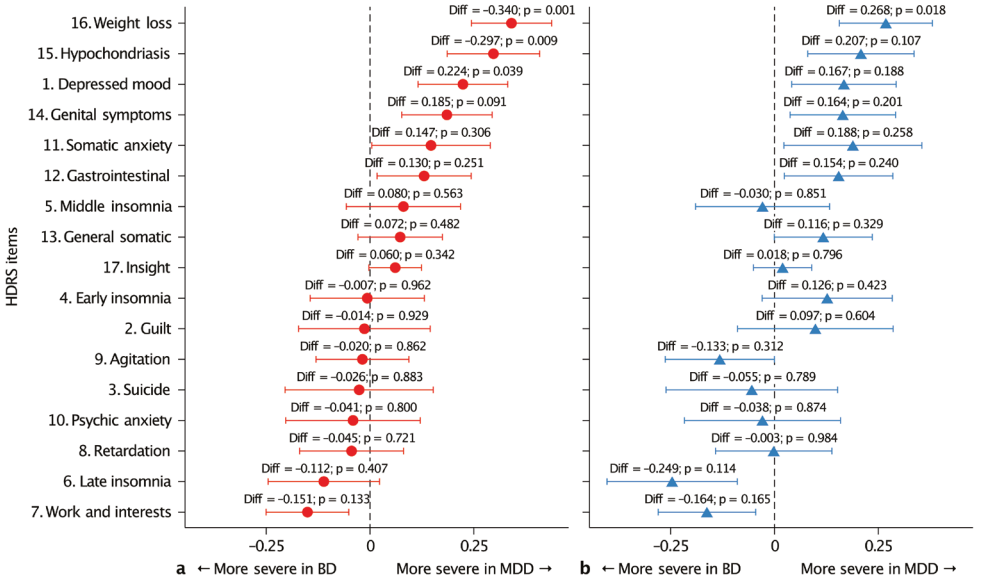


Figure. 1. Differences in individual symptoms at baseline in MDD and BD inpatients.
a. Unadjusted model. b Adjusted model. BD, bipolar depression; HDRS, Hamilton Depression Rating Scale; MDD, major depressive disorder.

Table 2 shows the use of psychotropic medication in the MDD and BD group at admission. In the MDD group, serotonin reuptake inhibitors were most commonly prescribed, followed by the selective serotonin-noradrenalin reuptake inhibitors and tricyclic antidepressants. Also in the BD group, SSRIs were most commonly prescribed, whereas lithium was the most frequently prescribed mood stabilizer.

Table 2: Psychotropic medication according to MDD and BD diagnoses at baseline

	Major Depressive Disorder (MDD, n=224)	Bipolar Depression (BD, n=52)	P-value*
Antidepressants:	78.1%	78.8%	0.91
- selective serotonin reuptake inhibitor (SSRI)	36.2 %	48.1 %	0.11
- serotonin–norepinephrine reuptake inhibitor (SNRI)	17.4 %	3.8 %	0.01
- tricyclic antidepressants (TCA)	23.7 %	17.3 %	0.32
Mood stabilizers:	5.4 %	48.1 %	< 0.001
- lithium	4.0 %	26.9 %	< 0.001
- lamotrigine	0.4 %	13.5 %	< 0.001
- sodium valproate	1.8 %	9.6 %	0.004
Antipsychotics:	42.4 %	53.8 %	0.14
- quetiapine	12.1 %	23.1 %	0.04
Benzodiazepines:	45.5 %	53.8 %	0.28

BD, bipolar depression; MDD, major depressive disorder

Concordance between the MINI and clinical diagnosis

All patients were diagnosed with a MINI-Plus mood disorder-either Mini-Plus MDD (*n* = 224) or MINI-Plus BD (*n* = 52), compared to a clinical diagnosis of MDD in 193 patients (69.9%) and BD in 55 patients (19.9%; Table 3). The remaining 10.2% of patients had either MINI-Plus MDD, diagnosed by clinicians as DSM-IV adjustment disorders (*n* = 23) or schizoaffective disorders (*n* = 2), or patients who had MINI-Plus BD clinically diagnosed as a DSM-IV adjustment disorder (*n* = 1) or schizoaffective disorders (*n* = 2). Clinical bipolar patients can be subdivided into BD Type I (*n* = 17), Type II (*n* = 31), or BD Not Otherwise Specified (NOS; *n* = 7). The agreement between a MINI and the clinical diagnosis of BD was moderate to strong with an efficiency of 89.5, PPV of 75, and NPV of 92.9; for unipolar depression, these values were comparable, 81.5, 81.7, and 80.8, respectively. The number of false positives was proportionally higher for MDD than for BD, resulting in a low specificity for MDD (50.6%) compared to BD (94.1%) (Table 3).

Table 3. Concordance between the MINI-Plus and clinical diagnosis in MDD and BD inpatients

	MDD	BD
Prevalence MINI-Plus (%)	81.2	18.8
Prevalence clinical diagnosis (%)	69.9	19.9
True positive (*n*)	183 (81.7%)	39 (75%)
False positive (*n*)	41	13
False negative (*n*)	10	16
True negative (*n*)	42	208
Area under the curve	0.73	0.83
Sensitivity (%)	94.8	70.9
Specificity (%)	50.6	94.1
Positive Predictive Value (%)	81.7	75.0
Negative Predictive Value (%)	80.8	92.9
Efficiency (n)	81.5	89.5

BD, bipolar depression; MDD: major depressive disorder, MINI-Plus, Mini-International Neuropsychiatric Interview Plus.

Time to response and remission

Table 4 shows the time to response and remission of the MINI-Plus MDD and Mini-Plus BD groups as well as the clinically diagnosed MDD and BD groups. Response was achieved in 60.7% of patients in the (MINI-Plus) MDD group and 63.5% in the (MINI-Plus) BD group; remission rates were 45.5% and 50.0%, respectively (Table 4).

Table 4. Response and remission according to MINI-Plus diagnoses in 276 in MDD and BD inpatients

	MDD	BD		*p* for interaction with time[a]
	No.(%) / HR	No.(%) / HR	*p* value	
MINI-Plus diagnoses:				
Response (*n* = 169 / 276; 61.2%)				
no./total no. (%)	136 / 224 (60.7%)	33 / 52 (63.5%)		
Crude	1.0 (ref.)	1.12 (0.76–1.64)	0.56	0.01
< 28 days[b]	1.0 (ref.)	0.49 (0.22–1.07)	0.07	
≥ 28 days[b]	1.0 (ref.)	1.68 (1.08–2.62)	0.02	
Adjusted[b,c]	1.0 (ref.)	1.23 (0.80-1.87)	0.34	0.04
< 28 days[b]	1.0 (ref.)	0.54 (0.23-1.23)	0.14	
≥ 28 days[b]	1.0 (ref.)	1.93 (1.16-3.20)	0.01	
Remission (*n* = 128 / 276; 46.4%)				
no./total no. (%)	102 / 224 (45.5%)	26 / 52 (50.0%)		
Crude[c]	1.0 (ref.)	1.14 (0.74–1.75)	0.56	0.69
< 28 days[b]	1.0 (ref.)	0.72 (0.28–1.85)	0.49	
≥ 28 days[b]	1.0 (ref.)	1.32 (0.81–2.14)	0.27	
Adjusted[b,c]	1.0 (ref.)	1.17 (0.73-1.89)	0.52	0.48
< 28 days[b]	1.0 (ref.)	0.73 (0.26-1.99)	0.53	
≥ 28 days[b]	1.0 (ref.)	1.40 (0.81-2.42)	0.23	
Clinical diagnoses:				
Response (*n* = 154 / 248; 62.1%)				
no./total no. (%)	114 / 193 (59.1%)	40 / 55 (72.7%)		
Crude	1.0 (ref.)	1.20 (0.81–1.77)	0.37	0.02
< 28 days[b]	1.0 (ref.)	0.80 (0.40–1.59)	0.52	
≥ 28 days[b]	1.0 (ref.)	1.64 (1.06–2.52)	0.03	
Adjusted[b,c]	1.0 (ref.)	1.40 (0.93-2.11)	0.10	0.07
< 28 days[b]	1.0 (ref.)	0.95 (0.44-2.03)	0.89	
≥ 28 days[b]	1.0 (ref.)	1.70 (1.04-2.79)	0.03	
Remission (*n* = 117 / 248; 47.2%)				
no./total no. (%)	87 / 193 (45.1%)	30 / 55 (54.5%)		
Crude[c]	1.0 (ref.)	1.18 (0.78–1.78)	0.45	0.91
< 28 days[b]	1.0 (ref.)	1.37 (0.61–3.07)	0.45	
≥ 28 days[b]	1.0 (ref.)	1.13 (0.67–1.89)	0.65	
Adjusted[b,c]	1.0 (ref.)	1.15 (0.72-1.86)	0.56	0.89
< 28 days[b]	1.0 (ref.)	1.83 (0.75-4.49)	0.19	
≥ 28 days[b]	1.0 (ref.)	1.02 (0.57-1.86)	0.94	

BD, bipolar depression; CI, confidence interval; HR, hazard ratio; MDD, major depressive disorder; MINI-Plus, Mini-International Neuropsychiatric Interview Plus. aInteraction with loge time (as a continuous variable). b Adjusted for sex, age (as a continuous variable), level of education, index episode, having ≥ 4 episodes, and duration of illness. c data are HRs with 95% CIs and p values, with the use of Cox proportional-hazards models.

Figures 2A and 2B show Kaplan-Meier curves for response and remission in the MINI-Plus MDD and BD groups. The Kaplan-Meier curve showed an initial advantage of MDD over BD for reaching response in the first 28 days, followed by a disadvantage (Fig 2A). Since the proportional hazard assumption failed (using the loglog graphical method; the *p* value for interaction with $\log_e$ time was 0.012), it would be misleading to report the difference between MDD and BD through the estimated HR, which assumes proportional hazards over time, because the difference depended on follow-up time [41].

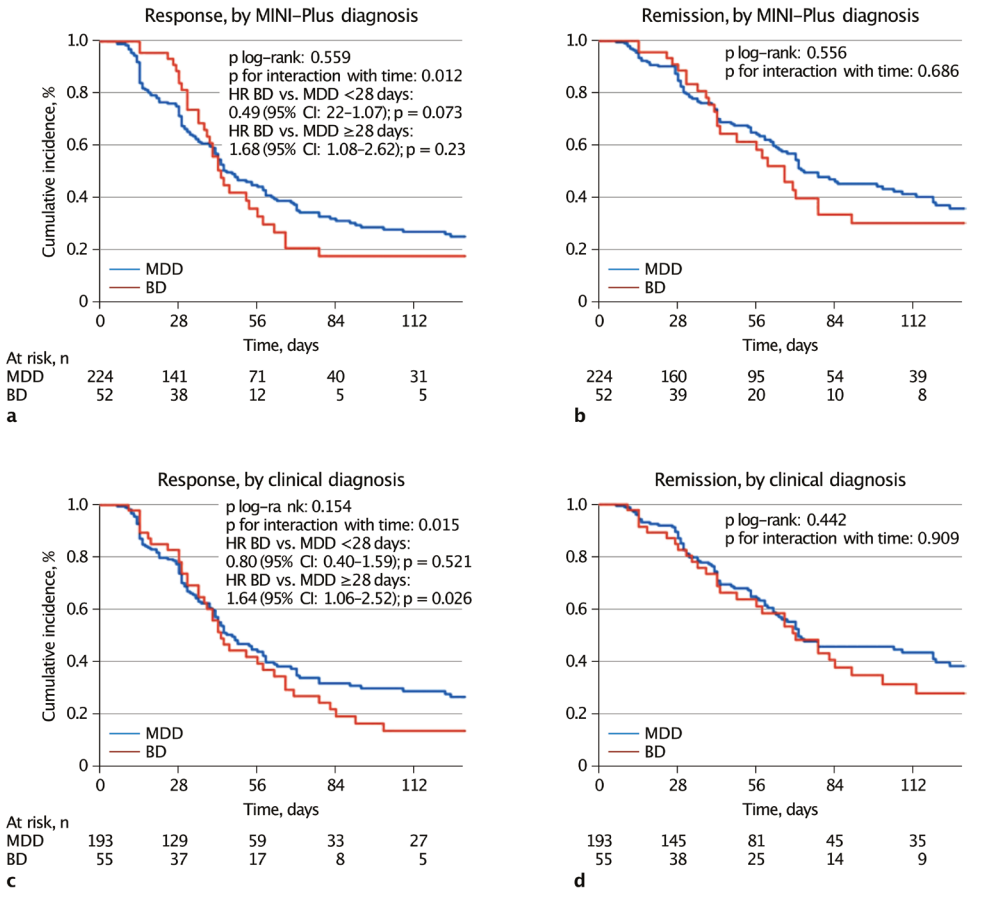


Figure 2. Kaplan-Meier curves for response and remission in MDD and BD inpatients.
a The Kaplan-Meier curve showed an initial advantage of MDD over BD for reaching response in the first 28 days, followed by a disadvantage. b Remission rates in the MDD and BD groups showed a rather similar course relative to each other. c, d Response and remission Kaplan-Meier curves by clinically diagnosed MDD and BD. BD, bipolar depression; CI, confidence interval; HR, hazard ratio; MDD, major depressive disorder; MINI-Plus, Mini-International Neuropsychiatric Interview Plus.

Adding a time-dependent covariate (i.e., $\log_e$ time) to the Cox proportional hazard model allowed us to analyze the change in the HR for response over time. Next, separate Cox regression analyses were performed in two time windows: before and after 28 days of admission, which on visual inspection seemed to be the "turning point" (i.e., when the HR changed course). MDD versus BD tended to have a higher rate of response during the first 4 weeks of admission with approaching statistical significance, HR = 0.54, 95% CI [0.23, 1.23], $p = 0.14$, but BD confirmed a significant long-term benefit on the likelihood of response, HR = 1.93, 95% CI [1.16, 3.20], p = 0.01 (Table 4). In contrast, remission rates in the MDD and BD groups showed a rather similar course relative to each other (Fig 2B).

Figures 2C and 2D show the response and remission Kaplan-Meier curves by clinically diagnosed MDD and BD groups. The BD group seemed to reach a faster response after 28 days, but the interaction term with $\log_e$ time in the crude analysis ($p = 0.02$) did not persist in the adjusted analysis ($p = .07$). No statistically significant interactions with time were found for time-to-remission for crude ($p = .91$) and for adjusted analyses ($p = 89$); see Table 4.

Discussion

In this study, we found that MDD and BD showed only minor differences in their symptom profiles at time of admission to the psychiatric hospital, with MDD patients showing more weight loss. The concordance indices between the MINI-Plus and clinical diagnoses were moderately strong, confirming the additive value of the MINI-Plus as a ROM diagnostic tool. Furthermore, BD patients showed a higher response rate during hospitalization compared to MDD patients, which is largely consistent with previous reports [20–23]. We found the same tendency for remission, though it was not statistically significant.

The prevalence of BD in patients with a depressive episode was 23.2% in our study, which is comparable to other prospective observational studies [42]. Consistent with previous studies [6,43-45], baseline depression severity (as assessed by the HDRS-17) was comparable in both groups. The few differences found in symptom profile between BD and MDD patients are partly in line with previous research that also found more weight loss [14] and more hypochondriasis [46] among MDD versus BD patients. Psychotropic medication, which is known to induce weight change, could only explain part of the difference in the multivariate analysis. Atypical depressive symptoms, which are usually more prevalent in BD [6–12], are not assessed by the HDRS-17, which limits its power to evaluate the symptom profile of BD. We could not replicate the higher occurrence of psychomotor retardation, which is usually more prevalent in MDD [6,9,13]. Also, presence of a MINI-Plus depression with melancholic features was more common in the MDD group than in the BD group; we

expected the opposite to be true based on previous research [6,9]. However, diagnosing melancholic features based on the MINI-Plus is not exempt from criticism and a more objective scale (e.g., the CORE scale [47]) would be more accurate.

As for the concordance of the MINI and clinical diagnoses, our findings are relevant for clinicians who need to differentiate between MDD and BD in a depressed inpatient, which is of clinical importance because treatment options differ significantly. Antidepressant monotherapy in BD is discouraged in current guidelines due to its low efficacy and even potential iatrogenic effects, particularly with TCA monotherapy, that is, switching to (hypo) mania or acceleration of mood cycles [48–50]. Although psychotherapeutic strategies in MDD and BD show some overlap, there are many psychotherapeutic and psychoeducational strategies that specifically target BD, such as Interpersonal Social Rhythm Therapy [49]. Since our sample consisted solely of inpatients with depressive symptomatology, our findings are difficult to compare directly to studies that investigated agreement between the MINI and clinical diagnoses in broader samples of psychiatric inpatients [51–53]. Nevertheless, the agreement indices of moderate strengths between the MINI and clinical diagnoses were comparable with those from Verhoeven et al. [32], who found an efficiency of 0.97, a PPV of 54.8%, and an NPV of 87.6% for BD (compared to 0.90, 0.75%, and 0.93% in our study). For unipolar depression, these values were 0.65, 54.8, and 87.6, respectively (compared to 0.82, 0.82, and 0.81). Therefore, we think that use of the MINI-Plus in ROM has added value for the clinician, not only because it assists the clinician in retrospectively assessing the presence of (hypo)mania and to ascertain either MDD or BD, but also to detect comorbid psychiatric diagnoses. Therefore, we think that use of the MINI in ROM has added value to the clinician, not only because it assists the clinician to retrospectively assess the presence of (hypo)mania and to ascertain either MDD or BD but also to detect comorbid psychiatric diagnoses.

To our knowledge, ours is the first prospective observational study examining short-term depressive symptom response in inpatients with MDD versus BD. Some previous prospective studies have shown that BD typically resolves faster than MDD [20,22,23]. Our finding that BD patients showed a higher response rate contrasts with the finding from the study by Furukuwa et al. [20] that showed no significant differences between MDD and BD in terms of time to recovery of depressive symptoms. In the latter study, however, sample size was smaller with 101 patients; the outcome measure (i.e., "recovery") was also defined differently, namely, as 2 consecutive months with no more than two mild depressive symptoms. Furthermore, Furukuwa et al.'s [20] population consisted solely of patients with an index episode. Although Kaplan-Meier curves were crossing, no analysis of the HR over time was performed.

Our study has several strengths. It featured a real-world naturalistic setting that included consecutive patients, having the advantage over randomized controlled trials of having less strict inclusion criteria with less selection bias, being more representative of psychiatric practice. Finally, we incorporated ROM assessments into clinical practice, implemented in short-term (2-week) intervals.

In addition to these strengths, we must also address some limitations. First, as a consequence of the naturalistic design, patients were treated with a variety of different combinations of psychotropic drugs that may have an influence on symptom profile and course. Moreover, we did not adjust for multiple testing, and therefore our findings need to be interpreted cautiously and to be replicated in independent samples. Second, sample sizes were moderate and the number of MDD patients was much larger than that of BD patients, which however reflects the real-world ratio of MDD and BD patients among consecutive inpatients. Third, the setting of a single tertiary psychiatric hospital may have limited the generalizability of our results. Fourth, the HDRS-17 may not be optimal to assess BD because it does not assess atypical depressive symptoms; however, it did facilitate the direct comparison to unipolar depression. Finally, regarding the agreement between the MINI-Plus and clinical diagnoses, there was a time lag between MINI-Plus diagnostics (at baseline) and clinical diagnostics (i.e., overall consensus of clinical diagnosis during hospitalization), which is an important limitation.

In conclusion, this study showed that ROM may help clinicians to diagnose patients who are admitted to a psychiatric setting and to monitor treatment response over time. Using ROM, we found that MDD patients exhibited more weight loss compared to their BD counterparts. Furthermore, although BD inpatients showed an initial slower rate of response compared to their MDD counterparts, more BD patients eventually reached response.

References

1. Forty L, Jones L, Jones I, Smith DJ, Caesar S, Fraser C, et al. Polarity at illness onset in bipolar I disorder and clinical course of illness. Bipolar Disord 2009;11:82–8.
2. Kupfer DJ, Frank E, Grochocinski VJ, Cluss PA, Houck PR, Stapf DA. Demographic and clinical characteristics of individuals in a bipolar disorder case registry. J Clin Psychiatry 2002;63:120–5.
3. Perlis RH, Delbello MP, Miyahara S, Wisniewski SR, Sachs GS, Nierenberg AA. Revisiting depressive-prone bipolar disorder: Polarity of initial mood episode and disease course among bipolar I systematic treatment enhancement program for bipolar disorder participants. Biol Psychiatry 2005;58:549–53.
4. Mitchell PB, Goodwin GM, Johnson GF, Hirschfeld RM a. Diagnostic guidelines for bipolar depression: a probabilistic approach. Bipolar Disord 2008;10:144–52.
5. Bowden CL. A different depression: Clinical distinctions between bipolar and unipolar depression. J Affect Disord 2005;84:117–25.
6. Mitchell PB, Wilhelm K, Parker G, Austin MP, Rutgers P, Malhi GS. The clinical features of bipolar depression: a comparison with matched major depressive disorder patients. J Clin Psychiatry 2001;62:212–216; quiz 217.
7. Serretti A, Mandelli L, Lattuada E, Cusin C, Smeraldi E. Clinical and demographic features of mood disorder subtypes. Psychiatry Res 2002;112:195–210.
8. Benazzi F. Clinical differences between bipolar II depression and unipolar major depressive disorder: lack of an effect of age. J Affect Disord 2003;75:191–5.
9. Benazzi F. Symptoms of depression as possible markers of bipolar II disorder. Prog Neuropsychopharmacol Biol Psychiatry 2006;30:47.
10. Andreasen NC, Grove WM, Coryell WH, Endicott J, Clayton PJ. Bipolar versus unipolar and primary versus secondary affective disorder: which diagnosis takes precedence? J Affect Disord n.d.;15:69–80.
11. Guze SB, Woodruff RA, Clayton PJ. The significance of psychotic affective disorders. Arch Gen Psychiatry 1975;32:1147–50.
12. Coryell W, Endicott J, Maser JD, Keller MB, Leon AC, Akiskal HS. Long-term stability of polarity distinctions in the affective disorders. Am J Psychiatry 1995;152:385–90. doi:10.1176/ajp.152.3.385.
13. Parker G, Roy K, Wilhelm K, Mitchell P, Hadzi-Pavlovic D. The nature of bipolar depression: implications for the definition of melancholia. J Affect Disord 2000;59:217–24.
14. Abrams R, Taylor MA. A comparison of unipolar and bipolar depressive illness. Am J Psychiatry 1980;137:1084–7.
15. Brockington IF, Altman E, Hillier V, Meltzer HY, Nand S. The clinical picture of bipolar affective disorder in its depressed phase. A report from London and Chicago. Br J Psychiatry 1982;141:558–62.
16. Katz MM, Robins E, Croughan J, Secunda S, Swann A. Behavioural measurement and drug response characteristics of unipolar and bipolar depression. Psychol Med 1982;12:25–36.

17. Beigel A, Murphy DL. Unipolar and bipolar affective illness. Differences in clinical characteristics accompanying depression. Arch Gen Psychiatry 1971;24:215–20.
18 Hegerl U, Bottner A-C, Holtschmidt-Täschner B, Born C, Seemüller F, Scheunemann W, et al. Onset of depressive episodes is faster in patients with bipolar versus unipolar depressive disorder: evidence from a retrospective comparative study. J Clin Psychiatry. 2008;69:1075–80.
19. Gassab L, Mechri A, Gaha L, Khiari G, Zaafrane F, Zougaghi L. [Bipolarity correlated factors in major depression: about 155 Tunisian inpatients]. Encephale n.d.;28:283–9.
20. Furukawa TA, Konno W, Morinobu S, Harai H, Kitamura T, Takahashi K. Course and outcome of depressive episodes: comparison between bipolar, unipolar and subthreshold depression. Psychiatry Res 2000;96:211–20.
21. Parker G, Fletcher K, Barrett M, Breakspear M, Rees A-M. Evaluating the first 1000 patients referred to a specialist depression clinic: a case for tertiary referral facilities. J Affect Disord 2011;131:52–8.
22. Winokur G, Coryell W, Keller M, Endicott J, Akiskal H. A prospective follow-up of patients with bipolar and primary unipolar affective disorder. Arch Gen Psychiatry 1993;50:457–65.
23 Angst J, Preisig M. Course of a clinical cohort of unipolar, bipolar and schizoaffective patients. Results of a prospective study from 1959 to 1985. Schweiz Arch Neurol Psychiatr 1995;146:5–16.
24. Kessing L V, Mortensen PB. Recovery from episodes during the course of affective disorder: a case-register study. Acta Psychiatr Scand 1999;100:279–87.
25. Goldberg JF, Harrow M. Consistency of remission and outcome in bipolar and unipolar mood disorders: A 10-year prospective follow-up. J Affect Disord 2004;81:123–31.
26. De Beurs E, Den Hollander-Gijsman ME, Van Rood YR, Van Der Wee NJA, Giltay EJ, Van Noorden MS, et al. Routine outcome monitoring in the Netherlands: Practical experiences with a web-based strategy for the assessment of treatment outcome in clinical practice. Clin Psychol Psychother 2011.
27. van Noorden M, van der Wee N, Zitman F, Giltay E, van Noorden MS. Routine outcome monitoring in psychiatric clinical practice: background, overview and implications for person-centered psychiatry. Eur J Pers Centered Healthc 2012;1:103– 11.
28. Washington AE, Lipstein SH. The Patient-Centered Outcomes Research Institute — Promoting Better Information, Decisions, and Health. N Engl J Med 2011;365:e31.
29. Wittchen HU. Reliability and validity studies of the WHO--Composite International Diagnostic Interview (CIDI): a critical review. J Psychiatr Res n.d.;28:57–84.
30. Sheehan D V., Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. 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. J Clin Psychiatry 1998;59:22–33.
31. Rettew DC, Lynch AD, Achenbach TM, Dumenci L, Ivanova MY. Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. Int J Methods Psychiatr Res 2009;18:169–84.

32. Verhoeven FEA, Swaab LSMA, Carlier IVE, van Hemert AM, Zitman FG, Ruhé HG, et al. Agreement between clinical and MINI diagnoses in outpatients with mood and anxiety disorders. J Affect Disord 2017;221:268–74.
33. van Vliet IM, de Beurs E. [The MINI-International Neuropsychiatric Interview. A brief structured diagnostic psychiatric interview for DSM-IV en ICD-10 psychiatric disorders]. Tijdschr Psychiatr 2007;49:393–7.
34. Hamilton M. A RATING SCALE FOR DEPRESSION. J Neurol Neurosurg Psychiatry 1960;23:56–62.
35. Trajković G, Starčević V, Latas M, Leštarević M, Ille T, Bukumirić Z, et al. Reliability of the Hamilton Rating Scale for Depression: a meta-analysis over a period of 49 years. Psychiatry Res 2011;189:1–9.
36. Beck AT, Steer RA BO. Beck Depression Inventory Manual, 2nd ed. Psychological Corporation, San Antonio, TX; n.d.
37. Wang Y-P, Gorenstein C. Psychometric properties of the Beck Depression Inventory- II: a comprehensive review. Rev Bras Psiquiatr 2013;35:416–31.
38. Sheehan D V, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. 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. J Clin Psychiatry 1998;59 Suppl 2:22-33;quiz 34-57.
39. Lecrubier Y, Sheehan D, Weiller E, Amorim P, Bonora I, Harnett Sheehan K, et al. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. Eur Psychiatry 1997;12:224–31.
40. Sheehan D, Lecrubier Y, Harnett Sheehan K, Janavs J, Weiller E, Keskiner A, et al. The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. Eur Psychiatry 1997;12:232–41.
41. Putter H, Sasako M, Hartgrink HH, van de Velde CJH, van Houwelingen JC. Longterm survival with non-proportional hazards: Results from the Dutch Gastric Cancer Trial. Stat Med 2005;24:2807–21.
42. Angst J, Azorin J-M, Bowden CL, Perugi G, Vieta E, Gamma A, et al. Prevalence and characteristics of undiagnosed bipolar disorders in patients with a major depressive episode: the BRIDGE study. Arch Gen Psychiatry 2011;68:791–8. doi:10.1001/archgenpsychiatry.2011.87.
43. Dorz S, Borgherini G, Conforti D, Scarso C, Magni G. Depression in Inpatients: Bipolar vs Unipolar. Psychol Rep 2003;92:1031–9. doi:10.2466/pr0.2003.92.3.1031.
44. Perlis RH, Brown E, Baker RW, Nierenberg AA. Clinical features of bipolar depression versus major depressive disorder in large multicenter trials. Am J Psychiatry 2006;163:225–31.
45. Ahearn EP, Carroll BJ. Short-term variability of mood ratings in unipolar and bipolar depressed patients. J Affect Disord 1996;36:107–15.
46. Sung G, Kim B-N, Lee E-H, Yu B-H, Hong KS, Kim J-H. Underestimating the severity of bipolar depression: a comparison of the Hamilton Depression Rating Scale items. J Affect Disord 2012;136:425–9.

47. Parker G, Hadzi-Pavlovic D. Development and Structure of the CORE System. In: Parker G, Hadzi-Pavlovic D, editors. Melancholia A Disord. Mov. Mood, Cambridge: Cambridge University Press; 1996, p. 82–129.
48. Kendall T, Morriss R, Mayo-Wilson E, Marcus E. Assessment and management of bipolar disorder: summary of updated NICE guidance. BMJ 2014;349:g5673–g5673.
49. Milev R V., Giacobbe P, Kennedy SH, Blumberger DM, Daskalakis ZJ, Downar J, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder. Can J Psychiatry 2016;61:561–75.
50. Goodwin GM, Haddad PM, Ferrier IN, Aronson JK, Barnes T, Cipriani A, et al. Evidence-based guidelines for treating bipolar disorder: Revised third edition Recommendations from the British Association for Psychopharmacology. J Psychopharmacol 2016;30:495–553.
51. Mordal J, Gundersen O, Bramness JG. Norwegian version of the Mini-International Neuropsychiatric Interview: feasibility, acceptability and test-retest reliability in an acute psychiatric ward. Eur Psychiatry 2010;25:172–7.
52. Bohnen EMA, De Winter RFP, Hoencamp E. Diagnostiek met MINI-plus in de acute psychiatrie. Tijdschr Psychiatr 2011;53:239–44.
53. Otsubo T, Tanaka K, Koda R, Shinoda J, Sano N, Tanaka S, et al. Reliability and validity of Japanese version of the Mini-International Neuropsychiatric Interview. Psychiatry Clin Neurosci 2005;59:517–26.

CHAPTER 3

Tryptophan catabolites in bipolar disorder: A meta-analysis

Hebbrecht K, Skorobogatov K, Giltay EJ, Coppens V, De Picker L, Morrens M.

Front Immunol. 2021 May 19;12:667179.

Abstract

Objective: Tryptophan catabolites (TRYCATs) are implicated in the pathophysiology of mood disorders by mediating immune-inflammation and neurodegenerative processes. We performed a meta-analysis of TRYCAT levels in bipolar disorder (BD) patients compared to healthy controls.

Methods: A systematic literature search in seven electronic databases (PubMed, Embase, Web of Science, Cochrane, Emcare, PsycINFO, Academic Search Premier) was conducted on TRYCAT levels in cerebrospinal fluid or peripheral blood according to the PRISMA statement. A minimum of three studies per TRYCAT was required for inclusion. Standardized mean differences (SMD) were computed using random effect models. Subgroup analyses were performed for BD patients in a different mood state (depressed, manic). The methodological quality of the studies was rated using the modified Newcastle-Ottawa Quality assessment Scale.

Results: Twenty-one eligible studies were identified. Peripheral levels of tryptophan (SMD = -0.44; $p < 0.001$), kynurenine (SMD = - 0.3; $p = 0.001$) and kynurenic acid (SMD = -.45; $p = < 0.001$) were lower in BD patients versus healthy controls. In the only three eligible studies investigating TRP in cerebrospinal fluid, tryptophan was not significantly different between BD and healthy controls. The methodological quality of the studies was moderate. Subgroup analyses revealed no significant difference in TRP and KYN values between manic and depressed BD patients, but these results were based on a limited number of studies.

Conclusion: The TRYCAT pathway appears to be downregulated in BD patients. There is a need for more and high-quality studies of peripheral and central TRYCAT levels, preferably using longitudinal designs

Introduction

Bipolar disorder (BD) is a chronic psychiatric disorder characterized by alternating periods of depression and abnormally elevated moods. BD is one of the leading causes of global disability, resulting in cognitive and functional decline and an increased mortality rate [1]. The pathophysiology of BD remains to be fully elucidated but accumulating evidence points towards a pathophysiological role of chronic low-grade inflammation [2].

The kynurenine pathway of tryptophan (TRP) degradation has been proposed as the missing link through which inflammation causes neurotoxicity and psychiatric symptoms. TRP is an essential amino acid and a precursor for serotonin or 5-hydroxytryptamine. In response to inflammation or psychosocial stress [3], TRP is primarily metabolized into kynurenine (KYN) following an upregulation of indoleamine 2,3-deoxygenase (IDO-1) and hereby leading to a reduction in availability of serotonin (for a graphical illustration of the KYN Pathway, see Figure 1). This depletion of serotonin has been assumed to play a major role in the pathophysiology of depression [5, 6]. More recent studies also point towards the imbalance supposedly neurotoxic [including 3-hydroxy kynurenine (3-HK) and quinolinic acid (QA)) and neuroprotective (kynurenic acid (KA)] TRP catabolites (TRYCAT) as a central mechanism in the pathophysiology of mood disorders [7, 8]. In patients with Major Depressive Disorder (MDD), a consistent increase in 3-HK and QA and a decrease in KA in blood and cerebrospinal fluid has been found [8, 9]. In BD patients, however, results have been more divergent and appear specific to the symptomatic state [10]. In depressed or euthymic BD patients, TRYCAT alterations seem to be similar to those in MDD [11–13]. In contrast, BD patients with a history of psychosis have shown elevated KA levels in cerebrospinal fluid (CSF) but not in the periphery, analogous to schizophrenia patients [13–16].

In the last decade, a growing number of studies in BD has been published and TRYCATs are represented as promising biomarkers related to BD [17]. However, studies show conflicting results and there is a great variation in methodological quality between studies, with a potential risk of bias as a consequence. Two previous meta-analyses synthesized the role of kynurenine metabolites in broad psychiatric disorders [18, 19]. Both included a limited number of studies in BD which investigated only a limited selection of TRYCATs (mostly TRP, KYN and/or KA) and the impact of mood state was not investigated. Arnone and colleagues [18] reported no significant differences in peripheral KYN or TRP values compared to controls, but only five studies were included and there was considerable heterogeneity among studies. The meta-analysis by Wang and Miller [19] found that CSF levels of KA were significantly increased in euthymic BD patients compared to healthy controls, but this finding were based upon two studies with partly overlapping samples [15, 16]. A third, recently published, meta-analysis summarized the results of studies on TRYCATs in BD, but they included only studies investigating TRYCAT levels in peripheral blood that

were published after 2006. Furthermore, they did not provide a critical evaluation of the study quality [20].

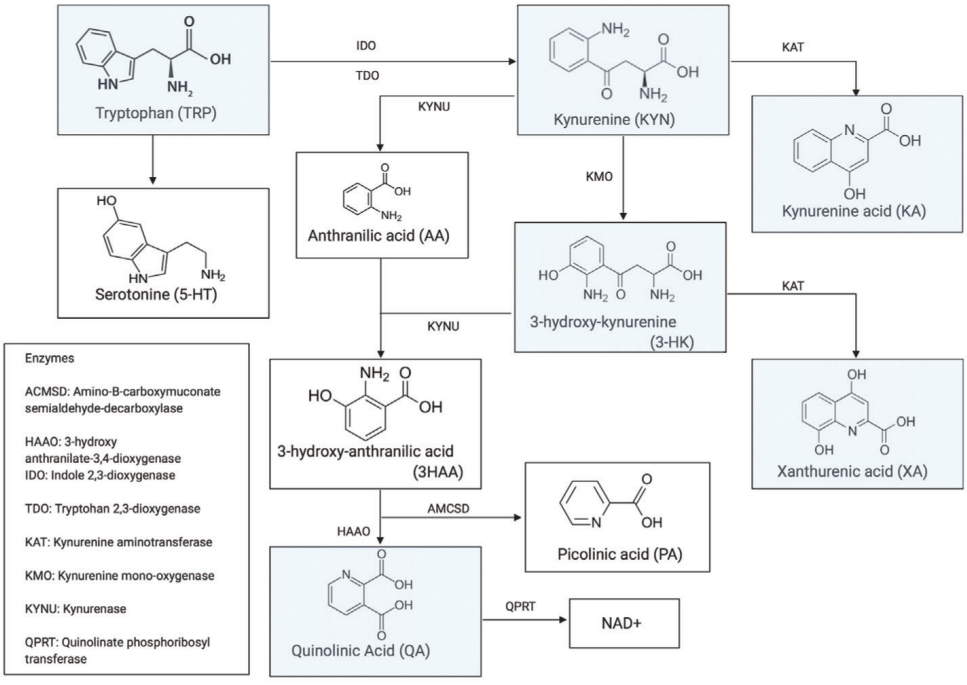


Figure 1. Kynurenine Pathway

The aim of this meta-analysis is to synthesize the available evidence on peripheral and central TRYCAT alterations in case- control studies of BD patients and to critically evaluate the quality of available studies. Furthermore, subgroup analyses were performed to separately investigate the differences in TRYCAT levels in manic (BD-M) and bipolar depressed (BD-D) patients.

Material and Methods

This meta-analysis was conducted and written according to the principles of the PRISMA-P (preferred reporting items for systematic review and meta-analysis protocols) guidelines [21].

Search Strategy

A search of seven electronic databases (PubMed, Embase, Web of Science, Cochrane, Emcare, PsycINFO, Academic Search Premier) was conducted for original papers on levels of TRYCATs (i.e. TRP, KYN, KA, 3-HK, QA) in BD patients. A medical librarian of the University of Leiden was involved in the establishment of the search strings (see Supplementary S1) and the literature search (last search: August, 19, 2020). Two authors (M.M. and K.H.) independently assessed studies for suitability for inclusion.

Inclusion criteria for eligible papers were: 1) English language papers published in peer-reviewed journals; 2) Case-control studies comparing BD patients (as confirmed by Research Diagnostic Criteria (RDC), DSM-(III, III-R, IV, IV-TR) or ICD-(9 or 10) to healthy controls, 3) assessment of at least 1 TRYCAT metabolite in peripheral blood, CSF or postmortem tissue. In case of sample overlap between studies (as indicated by the authors), only the largest study was included in the current meta-analysis, in order to avoid double counting. Only baseline data were included from longitudinal study articles.

Quality Assessment

Two researchers (KH and KS) independently assessed the risk of bias and methodological quality of the included studies using a modified version of the Newcastle-Ottawa Quality assessment Scale for case-control studies [22]. Following assessments were added to the original scale: an evaluation of the sample size (i.e. a required sample size of minimum twenty patients), assessment of outcome consisting of an evaluation of the completeness of TRYCAT description on the one hand as well as lab procedures (including blinding) in order to guarantee reproducibility on the other hand) and an assessment of statistical reporting. Studies could obtain up to ten stars on three overall quality domains (i.e. selection, comparability, and outcome).

Data Synthesis and Analysis

Demographic variables (age and gender), clinical assessments (mood state and symptom severity scores), and TRYCAT metabolite levels (means and standard deviations) were extracted from each study. Authors were contacted for additional information when data could not be extracted from the paper; this was received from four papers [13, 23–25]. The Review Manager 5.3 (RevMan 5.3) computer program was used for performing the primary meta-analysis and subgroup analyses. The primary outcome measure was the standardized mean difference (SMD) in random effect models, represented in forest plot graphs (95% confidence interval). The presence of heterogeneity was assessed using Chi^2 and its magnitude using I^2 statistics. Potential effect modification by gender, age, and publication year was investigated by performing metaregression analyses (Knapp-Hartung method, maximum likelihood) [26] in Comprehensive Meta-Analysis version 3

(CMA v3). For analyses with ten or more available studies, funnel plots and Egger's tests were used to assess the presence of publication bias.

Subgroup analyses were performed to investigate the difference in TRYCAT levels for BP patients in a different mood state (depressed, manic). A minimum of three studies per subgroup was required in order to perform a subgroup analysis for each TRYCAT. A depressed state was defined as a major depressive episode as diagnosed by the RDC, DSM-(III, III-R, IV, IV-TR) or ICD-(9 or 10) criteria and/or defined as a minimum threshold of 17 or 18 on the Hamilton Rating Scale for Depression (HRSD-17) or 20 on the HRSD-24 [27]. A manic state was defined as fulfilling the criteria of the RDC, DSM-(III, III-R, IV, IV-TR) or ICD-(9 or 10) ICD-10) criteria and/or as having a minimum threshold of 13 on the Young Mania Rating Scale (YMRS) [28], the most frequently used scale for assessment of manic symptoms. By means of a supplementary analysis, subgroup analyses were also performed to investigate the differences in effect size between high and low quality studies. The significance level was set at $p < 0.05$, the Benjamin Hochberg procedure was applied for controlling false discovery rates (FDR) in meta-regression analyses.

Results

Study Selection

The search strategy resulted in 903 hits and after deduplication 438 remained that were screened for relevance based on title and abstract. A final of 47 papers were read in full, of which 26 were excluded. The PRISMA Flow Diagram in Figure 2 depicts the number of in- and excluded articles from each stage of screening. Four studies investigated TRYCATs in CSF [15, 29–31], sixteen in serum or plasma and one both in CSF and serum [13]. Only one post-mortem study met inclusion criteria [32], but this study was excluded due to inadequate reporting. Of the twenty-one included papers in the meta-analysis, twelve had a cross-sectional design; nine a longitudinal design.

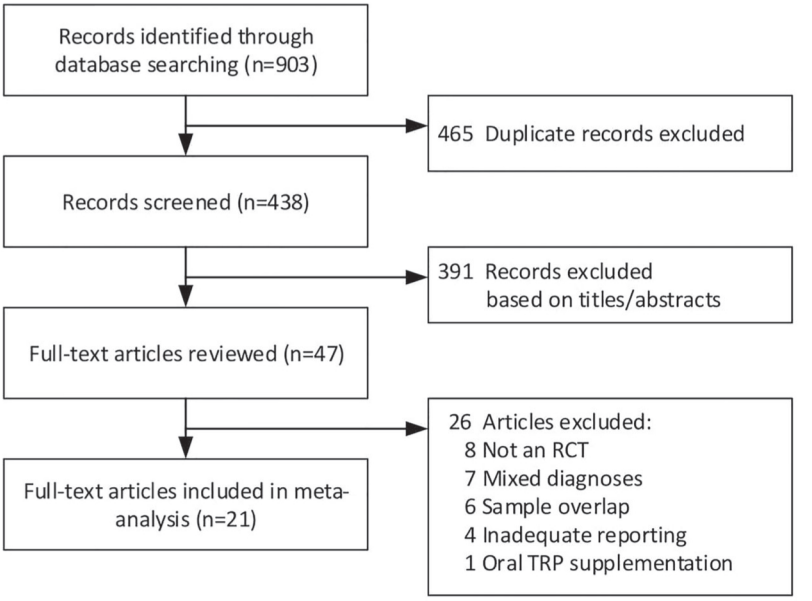


Figure 2. Flowchart

Table 1 presents the characteristics of the included studies. The analysis of TRP in CSF and five TRYCATs (TRP, KYN, KA, 3-HK, QA) in peripheral blood were included in the meta-analyses based on the minimal requirement of three studies for each meta-analysis. Two CSF studies included both BD-D and BD-M patients [29, 30], one solely BD-M [31] and one solely euthymic BD patients [15]. Eight serum/plasma studies included only BD-D [12, 23, 33, 35–37, 41, 42], one only BD-M [14], one only euthymic BD [39], two both BD-D and BD-M patients [24, 40], one study both BDM and euthymic BD [34] and two studies BD-D, BD-M and BD-Mixed patients [10, 17]. In the two remaining studies the mood state of BD patients was not specified [13, 38].

Table 1: Characteristics of included studies

Author, Year	No. of participants (P/HC)	Mean Age (SD) (P/HC)	%male (P/HC)	D/M/E	Sample Type	Longitudinal (yes/no)	TRP	KYN	3-HK	QA	KA
Ashcroft, 1973	13/26	NA	NA	D(6) M(7)	CSF	Yes	x				
Coppen, 1972	13/14	50.6 (4.09)/ NA		D(10), M(3)	CSF	No	x				
Gerner 1984	13/37	NA	NA	M(13)	CSF	No	x				
Olsson, 2012	55/23	39(14)/ 33.1(6.9)	38/100	E(55)	CSF	No					x
Sellgren, 2019	93/113	NA	NA	NA	CSF	No					x
Total number of studies in CSF tissue (n = 5)							**3**				**2**
Chiaroni, 1990	18/33	57.9(17.6)/ 39.5 (8.4)	36/42	D (18)	Plasma	Yes	x				
Hoekstra, 2006	32/20	47.2(14.6)/ 50.1 (13.5)	63/70	M (20) E (12)	Plasma	Yes	x				
Liu, 2018	20/23	30.45 (5.09)/ 29.3 (5.89)	40/52	D (20)	Plasma	No	x	x		x	x
Moller, 1979	18/25	39.4 (14.1)/ 41.7 (12.4)	0/0	D (18)	Plasma	No	x				
Mukherjee, 2018	21/28	36.1 (11.3)/ 31.6 (10.3)	52/43	D (9), M (1), Mixed (10)*	Plasma	Yes	x	x			
Murata, 2020	43/26	42.3 (12.3)/ 39.6 (13.5)	48/52	D (48)	Serum	Yes	x	x			
Myint, 2007	39/80	37.6 (11.6)/ 39.1 (8.8)	38/50	M (39)	Plasma	Yes	x	x			x
Olajossy, 2017	11/48	44.7 (13.8) /35 (NR)	43/48	D (11)	Serum	Yes					x
Pan, 2018	30/40	35.8 (10.7)/ 36.8 (8.8)	43/55	NR	Plasma	No		x			

Platzer, 2017	68/93	44.9 (14)/ 38.9 (16.2)	62/39	E (68)	Serum	No	x	x	x		x
Poletti, 2018	22/15	46.5 (13.7)/ 27.2 (8.3)	36/43	D (22)	Serum	No	x	x	x		x
Poletti, 2019	72/36	48.0 (13.7)/ 43.9 (12.3)	40/36	D (55), M (17)	Plasma	No	x	x			
Savitz, 2015	63/48	38.8 (11.1)/ 32.6 (10.4)	19/40	D (63)	Serum	No	x	x	x	x	x
Sellgren, 2019 (plasma)	163/114	34 (NR) / 35 (NR)	39/46	NR	Plasma	No					x
Van den Ameele 2020	65/34	43.1 (11.2)/ 42.7 (11.6)	42/46	D (35), M(32)	Plasma	Yes	x	x	x	x	x
Wurfel, 2017	53/92	40.2 (11.0)/ 32.3 (10.4)	30/36	D (15), M (25), Mixed (10)**	Serum	No	x	x	x	x	x
Zhou, 2018	16/6	37.8 (12.9)/ 31.6 (10.7)	69/58	D (16)	Serum	Yes	x	x			x
Total number of studies in peripheral blood tissue (n = 16)							14	12	5	4	10

*: one patient did not have CARS-M data

**: affective state missing from three patients

Abbreviations. P= Patients; HC = healthy controls; D = Depressed; M = Manic; E = Euthymic; TRP = Tryptophan; KYN = kynurenine; 3-HK = 3-hydroxykynurenine; QA = Quinolinic acid; KA = Kynurenic Acid; SMD = Standard Mean Difference; NA = not applicable

Quality Assessment

The results of the quality assessment can be found in Table 2. The quality analysis showed an overall moderate methodological quality with 12 studies (57%) scoring half of the maximum score or more (5/10 or more). Eight studies (38%) had a sample size of less than twenty patients [12, 23, 29–31, 33, 36, 42]. Only five studies recruited a matched control sample [13, 23, 24, 34, 38] and all but one study [10] reported unadjusted mean TRYCAT levels. Four studies reported that the laboratory technicians were blind for diagnose status [10, 14, 31, 41].

Table 2: Psychotropic medication according to MDD and BD diagnoses at baseline

	Major Depressive Disorder (MDD, n=224)	Bipolar Depression (BD, n=52)	P-value*
Antidepressants:	78.1%	78.8%	0.91
- selective serotonin reuptake inhibitor (SSRI)	36.2 %	48.1 %	0.11
- serotonin–norepinephrine reuptake inhibitor (SNRI)	17.4 %	3.8 %	0.01
- tricyclic antidepressants (TCA)	23.7 %	17.3 %	0.32
Mood stabilizers:	5.4 %	48.1 %	< 0.001
- lithium	4.0 %	26.9 %	< 0.001
- lamotrigine	0.4 %	13.5 %	< 0.001
- sodium valproate	1.8 %	9.6 %	0.004
Antipsychotics:	42.4 %	53.8 %	0.14
- quetiapine	12.1 %	23.1 %	0.04
Benzodiazepines:	45.5 %	53.8 %	0.28

BD, bipolar depression; MDD, major depressive disorder

Central Levels of Kynurenine Metabolites

CSF levels of TRP did not significantly differ from healthy controls ($n_{studies} = 3$, $n_{patients} = 39$, SMD = - 0.43, z = 0.86, $p = 0.39$). There was considerable inter-study heterogeneity (I^2: 83%, see Supplementary Figure 4). Only two studies investigated KA in CSF in BD. No CSF studies were found for KYN, 3-HK and QA in BD. Consequently, these four TRYCATs were not included in the meta-analysis.

Peripheral Levels of Kynurenine Metabolites

Peripheral blood levels of TRP, KYN and KA were significantly lower in BD compared to healthy controls (TRP: $n_{studies} = 14$, $n_{patients} = 552$, SMD = -0.44, z = 4.94, $p < 0.001$; KYN: $n_{studies} = 12$, $n_{patients} = 514$, SMD = - 0.30, z = 3.21, $p = 0.001$; KA: $n_{studies} = 10$, $n_{patients} = 522$, SMD = - 0.45, z = 3.98, $p < .001$). Peripheral QA and 3-HK concentrations did not differ significantly between BD and healthy controls (QA: $n_{studies} = 4$, $n_{patients} = 203$, SMD = - 0.31,

$z = 1.37, p = 0.17$; 3-HK: $n_{studies} = 5$, $n_{patients} = 273$, SMD = - 0.78, $z = 0.54, p = 0.59$). Inter-study heterogeneity was present for all TRYCATs with I^2 ranging from 46 to 77%.

Publication Bias

Funnel plots (of metabolites with a minimum of 10 available studies; TRP, KYN, KA in peripheral blood) are presented in Supplementary Figures 1 to 3. The funnel plot of KA shows a significant asymmetry, confirmed by the Egger's test (shown in Table 3), which potentially indicates a publication bias in favor of research reporting lower KA levels in BD.

Table 3. Concordance between the MINI-Plus and clinical diagnosis in MDD and BD inpatients

	MDD	BD
Prevalence MINI-Plus (%)	81.2	18.8
Prevalence clinical diagnosis (%)	69.9	19.9
True positive (*n*)	183 (81.7%)	39 (75%)
False positive (*n*)	41	13
False negative (*n*)	10	16
True negative (*n*)	42	208
Area under the curve	0.73	0.83
Sensitivity (%)	94.8	70.9
Specificity (%)	50.6	94.1
Positive Predictive Value (%)	81.7	75.0
Negative Predictive Value (%)	80.8	92.9
Efficiency (n)	81.5	89.5

BD, bipolar depression; MDD: major depressive disorder, MINI-Plus, Mini-International Neuropsychiatric Interview Plus.

Subgroup Analyses and Meta-Regression

Subgroup analyses in euthymic patients could not be reliably performed due to the scarcity of such studies, as there were only three studies including euthymic BD patients, of which one presented CSF levels. Subgroup effect by either depressed or manic mood state for TRP and KYN did not show effect modification (Chi^2 test for subgroup differences were not significant, see Supplementary Figures 11–12). Subgroup analyses for KA, 3-HK and QA could not be performed since the minimum criterion of three studies in each subgroup was not fulfilled. The pooled effect estimate for TRP in the BD-M subgroup was slightly larger than that of the BD-D subgroup (BD-M: SMD = - 0.52; $z = 2.32$; $p = 0.02$; BD-D: SMD = - 0.43; $z = 2.96$; $p = 0.003$). The pooled effect estimates for KYN in BD-D and BD-M groups were comparable (BD-M: SMD = - 0.27; $z = 1.98$; $p = 0.05$; BD-D: SMD = - 0.38; $z = 2.8$; $p = 0.005$). Considerate within-subgroup heterogeneity remained, indicating that other unidentified factors likely affect TRYCAT levels in BD patients. By means of a

supplementary analysis (Supplementary Table 1), we performed a subgroup analysis comparing effect sizes in high and low quality studies and this indicated a significant subgroup effect for KYN ($p = 0.04$) and KA ($p = 0.04$) with low quality studies showing larger effect sizes compared to high quality studies. As demonstrated in the meta-regression analyses (see Supplementary Table 2), there was no effect modification for TRP, KYN and KA by age. The gender of the control group appeared to be a significant moderator of the effect in the studies comparing KA in BD and controls, yet this was no longer significant after correcting for false discovery rates. Metaregressions could not be performed for 3-HK and QA due to the low number of studies (n = 5 and n = 4 respectively).

Discussion

This meta-analysis summarizes the available evidence on a wide range of TRYCAT metabolites, representative for the whole kynurenine pathway, in BD patients compared to healthy controls. Patients with BD showed lower peripheral levels of TRP, KYN and KA compared to healthy controls. The levels of 3-HK and QA were not significantly different between healthy controls and BD. CSF levels of TRP showed no significant difference between BD and healthy controls, but this finding was based on only three studies.

Our results confirm that BD is associated with alterations in TRYCATs. However, these findings do not entirely correspond to the theoretically proposed hypotheses to explain the relationship between inflammation, kynurenine metabolism and BD. TRYCATs are assumed to act as inflammatory mediators and to cause neurodegeneration through neurotoxic effects [43], but the exact pathophysiological mechanism how TRYCATS influence BD symptoms and course remain unclear. The lower TRP levels in peripheral blood are consistent with the inefficient serotonin turnover in BD [14, 17], but our findings are not consistent with the theoretical hypothesis of an increased TRP breakdown, under low-grade inflammatory conditions [11], which would be expected to result in elevated KYN and KA levels. A plausible explanation for this inconsistency may be that a proposed microglial branch upregulation results in a reduced shunt towards the astrocytic branch, resulting in lower KYNA levels [44].

Our findings are in line with a recent meta-analysis by our group on TRYCAT alterations in schizophrenia spectrum disorder (SSD) which showed a partial downregulation of the kynurenine pathway (significantly lower levels of peripheral TRP in all SSD patients but especially in acute psychotic, younger patients and of peripheral KA and QA in symptomatic and/or older SSD patients [4]. Accumulating evidence shows that acute psychotic exacerbations are associated with different immunological alterations than non-acute states [45, 46] and our group previously hypothesized differences in state (i.e.

emerging during acute exacerbations) and trait immune markers (i.e. relatively unaltered throughout the disorder) in SSD [47], which could also be the case in BD.

However, it should be noted that peripheral, rather than central TRYCAT metabolites have been measured in most studies. An important question is to what extent CSF and plasma TRYCAT levels are correlated and how they differentially influence the pathophysiology of BD. TRP, KYN, and 3-HK easily cross the blood brain barrier by active transport, but the brain uptake of QA and KA is limited to passive diffusion due to their polarity [48]. Sellgren et al.[13] have demonstrated that peripheral KA levels do not mirror central levels in a large sample of BD and healthy controls. But other studies did show a correlation between QA and KA levels in serum and CSF in depressed patients with proven signs of inflammation levels [49,50]. A secondary issue concerns the binding capacity of TRP, KYN and KA to plasma proteins, such as albumin, but the exact result on peripheral values and blood-brain transport remains unclear [48]. Third, the peripheral kynurenine pathway is regulated by immune markers, steroids and growth factors [51–53] which can also potentially affect peripheral levels.

All analyses of studies investigating the TRYCAT levels in peripheral blood showed substantial between-study heterogeneity, with effect sizes varying noticeably between studies. This suggests that a number of confounders and study-specific variables contribute to the effect size and, consequently, to the divergence in study results. We investigated the role of mood state (manic or depressed state) in subgroup analyses but this did not explain a significant proportion of the between-study variance. In a further attempt to reveal study-specific characteristics related with heterogeneity, meta-regression analyses were performed but these revealed no significant associations between TRYCAT levels and variables such as age, gender and publication year. It should be emphasized that other, not-investigated, factors could play a role in this heterogeneity. We can broadly categorize these factors into three domains: methodological, clinical and conceptual issues. Apart from differences in methodological quality between studies, differences in lab techniques could also lead to heterogeneous results. Although Liquid-Chromatography Mass Spectrometry is currently considered as golden standard and consequently the most commonly used method, other techniques have been used in studies such as High-Pressure Liquid Chromatography and Atomic Absorption Spectrophotometry). Moreover, some TRYCAT metabolites (such as QA) have extremely low concentrations in peripheral blood tissue which tend to border the limits of the detection range of most of these methods, which may greatly affect reliability of some of these assessments. Several clinical factors are assumed to influence TRYCAT levels, the most of which is the use of psychotropic drugs. Several studies have demonstrated a moderating effect of anticonvulsants (e.g. valproate) on TRYCAT levels [24, 34] but there is a lack of large-scale studies. Moreover, age and duration of illness may similarly have an effect on TRYCAT changes, although the limited amount of

available studies do not allow for proper analyses of these effects. Lastly, between-study heterogeneity could be a reflection of underlying genetic, phenotypical or diagnostic diversity of BD patients included in different studies [54]. However, this heterogeneity, which may translate in differential impact on the TRYCAT pathway, has never been investigated in BD patient groups.

To our knowledge, this meta-analysis provides the most extensive summary of all available studies on a wide range of TRYCAT levels measured in CSF or serum/plasma in BD patients published to date. Compared to previously published meta-analyses [18, 19], we performed a broader literature search and provided a more complete analysis of the data by contacting authors for additional data on TRYCAT levels of BD subgroups. Other strengths of our study are the critical quality assessment of the included studies and the separate analysis of TRYCAT alterations in BD patients in a different mood (manic, depressed) resulting in a more nuanced picture of TRYCAT alterations in BD and adding evidence to the discussion on whether TRYCAT alterations should be considered as state or trait dependent changes. However, our results need to be interpreted in view of some limitations. Some analyses included only a small number of studies and the methodological quality of some studies was insufficient. The interpretation of our results is further limited by the differential use of psychopharmacological treatments between patients within and between studies as these are known to have a confounding influence on inflammatory mediators. The majority of the individual studies did not adjust the analysis for important confounders, such as age, gender, smoking status, duration of BD, (doses of) psychotropics, and symptom severity.

Supplementary Material

The Supplementary Material can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2021.667179/full#supplementary-material

References

1. Grande I, Berk M, Birmaher B, Vieta E. Bipolar Disorder. Lancet (2016) 387:1561–72. doi: 10.1016/S0140-67361500241-X
2. Drexhage RC, Hoogenboezem TH, Versnel MA, Berghout A, Nolen WA, Drexhage HA. The Activation of Monocyte and T Cell Networks in Patients With Bipolar Disorder. Brain Behav Immun (2011) 25:1206–13. doi: 10.1016/j.bbi.2011.03.013
3. Heisler JM, O'Connor JC. Indoleamine 2,3-Dioxygenase-Dependent Neurotoxic Kynurenine Metabolism Mediates Inflammation-Induced Deficit in Recognition Memory. Brain Behav Immun (2015) 50:115–24. doi: 10.1016/j.bbi.2015.06.022
4. Morrens M, De Picker L, Kampen JK, Coppens V. Blood-Based Kynurenine Pathway Alterations in Schizophrenia Spectrum Disorders: A Meta-Analysis. Schizophr Res (2020) 223:43–52. doi: 10.1016/j.schres.2020.09.007
5. Dantzer R. Role of the Kynurenine Metabolism Pathway in Inflammation-Induced Depression: Preclinical Approaches. Curr Top Behav Neurosci (2017) 31:117–38. doi: 10.1007/7854_2016_6
6. Widner B, Laich A, Sperner-Unterweger B, Ledochowski M, Fuchs D. Neopterin Production, Tryptophan Degradation, and Mental Depression -What is the Link? Brain Behav Immun (2002) 16(5):590–5. doi: 10.1016/S0889-15910200006-5
7. Myint AM, Kim YK. Network Beyond IDO in Psychiatric Disorders: Revisiting Neurodegeneration Hypothesis. Prog Neuropsychopharmacol Biol Psychiatry (2014) 48:304–13. doi: 10.1016/j.pnpbp.2013.08.008
8. Savitz J. Role of Kynurenine Metabolism Pathway Activation in Major Depressive Disorders. Curr Top Behav Neurosci (2017) 31:249–68. doi: 10.1007/7854_2016_12
9. Myint AM, Kim YK, Verkerk R, Scharpé S, Steinbusch H, Leonard B. Kynurenine Pathway in Major Depression: Evidence of Impaired Neuroprotection. J Affect Disord (2007b) 98:143–51. doi: 10.1016/j.jad.2006.07.013
10. Wurfel BE, Drevets WC, Bliss SA, McMillin JR, Suzuki H, Ford BN, et al. Serum Kynurenic Acid is Reduced in Affective Psychosis. Transl Psychiatry (2017) 7(5):e1115. doi: 10.1038/tp.2017.88
11. Birner A, Platzer M, Bengesser SA, Dalkner N, Fellendorf FT, Queissner R, et al. Increased Breakdown of Kynurenine Towards its Neurotoxic Branch in Bipolar Disorder. PLoS One (2017) 12:1–14. doi: 10.1371/journal.pone.0172699
12. Poletti S, Myint AM, Schüetze G, Bollettini I, Mazza E, Grillitsch D, et al. Kynurenine Pathway and White Matter Microstructure in Bipolar Disorder. Eur Arch Psychiatry Clin Neurosci (2018) 268:157–68. doi: 10.1007/s00406-016-0731-4
13. Sellgren CM, Gracias J, Jungholm O, Perlis RH, Engberg G, Schwieler L, et al. Peripheral and Central Levels of Kynurenic Acid in Bipolar Disorder Subjects and Healthy Controls. Transl Psychiatry (2019) 9(1):37. doi: 10.1038/s41398- 019-0378-9
14. Myint AM, Kim YK, Verkerk R, Park SH, Scharpé S, Steinbusch HWM, et al. Tryptophan Breakdown Pathway in Bipolar Mania. J Affect Disord (2007a) 102:65–72. doi: 10.1016/j.jad.2006.12.008

15. Olsson SK, Sellgren C, Engberg G, Landén M, Erhardt S. Cerebrospinal Fluid Kynurenic Acid is Associated With Manic and Psychotic Features in Patients With Bipolar I Disorder. Bipolar Disord (2012) 14:719–26. doi: 10.1111/ bdi.12009
16. Olsson SK, Samuelsson M, Saetre P, Lindström L, Jönsson EG, Nordin C, et al. Elevated Levels of Kynurenic Acid in the Cerebrospinal Fluid of Patients With Bipolar Disorder. J Psychiatry Neurosci (2010) 35:195–9. doi: 10.1503/jpn.090180
17. Mukherjee D, Krishnamurthy VB, Millett CE, Reider A, Can A, Groer M, et al. Total Sleep Time and Kynurenine Metabolism Associated With Mood Symptom Severity in Bipolar Disorder. Bipolar Disord (2018) 20:27–34. doi: 10.1111/bdi.12529
18. Arnone D, Saraykar S, Salem H, Teixeira AL, Dantzer R, Selvaraj S. Role of Kynurenine Pathway and its Metabolites in Mood Disorders: A Systematic Review and Meta-Analysis of Clinical Studies. Neurosci Biobehav Rev (2018) 92:477–85. doi: 10.1016/j.neubiorev.2018.05.031
19. Wang AK, Miller BJ. Meta-Analysis of Cerebrospinal Fluid Cytokine and Tryptophan Catabolite Alterations in Psychiatric Patients: Comparisons Between Schizophrenia, Bipolar Disorder, and Depression. Schizophr Bull (2018) 44:75–83. doi: 10.1093/schbul/sbx035
20. Bartoli F, Misiak B, Callovini T, Cavaleri D, Cioni RM, Crocamo C, et al. The Kynurenine Pathway in Bipolar Disorder: A Meta-Analysis on the Peripheral Blood Levels of Tryptophan and Related Metabolites. Mol Psychiatry (2020). doi: 10.1038/s41380-020-00913-1
21. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols PRISMA-P 2015 Statement. Rev Esp Nutr Humana y Diet (2016) 20:148–60. doi: 10.1186/2046-4053-4-1
22. Peterson J, Welch V, Losos M, Shea B, O' Connell D, Tugwell P, et al. The Newcastle-Ottawa Scale NOS for Assessing the Quality of Nonrandomised Studies in Meta-Analyses. Ottawa: Ottawa Hospital Research Institute (2011).
23. Olajossy M, Olajossy B, Wnuk S, Potembska E, Urbańska E. Blood Serum Concentrations of Kynurenic Acid in Patients Diagnosed With Recurrent Depressive Disorder, Depression in Bipolar Disorder, and Schizoaffective Disorder Treated With Electroconvulsive Therapy. Psychiatr Pol (2017) 51:455–68. doi: 10.12740/pp/61584
24. van den Ameele S, van Nuijs AL, Lai FY, Schuermans J, Verkerk R, van Diermen L, et al. A Mood State-Specific Interaction Between Kynurenine Metabolism and Inflammation is Present in Bipolar Disorder. Bipolar Disord (2020) 22:59–69. doi: 10.1111/bdi.12814
25. Zhou Y, Zheng W, Liu W, Wang C, Zhan Y, Li H, et al. Cross-Sectional Relationship Between Kynurenine Pathway Metabolites and Cognitive Function in Major Depressive Disorder. Psychoneuroendocrinology (2019) 101:72–9. doi: 10.1016/j.psyneuen.2018.11.001
26. Tipton E, Pustejovsky JE, Ahmadi H. Current Practices in Meta-Regression in Psychology, Education, and Medicine. Res Synth Methods (2019) 10:180–94. doi: 10.1002/jrsm.1339
27. Hamilton M. A Rating Scale for Depression. J Neurol Neurosurg Psychiatry (1960) 23:56–62. doi: 10.1136/JNNP.23.1.56
28. Young RC, Biggs JT, Ziegler VE, Meyer DA. A Rating Scale for Mania: Reliability,Validity and Sensitivity. Br J Psychiatry J Ment Sci (1978) 133:429–35. doi: 10.1192/bjp.133.5.429

29. Ashcroft GW, Blackburn IM, Eccleston D, Glen AI, Hartley W, Kinloch NE, et al. Changes on Recovery in the Concentrations of Tryptophan and the Biogenic Amine Metabolites in the Cerebrospinal Fluid of Patients With Affective Illness. Psychol Med (1973) 3:319–25. doi: 10.1017/s0033291700049606
30. Coppen A, Brooksbank BW, Peet M. Tryptophan Concentration in the Cerebrospinal Fluid of Depressive Patients. Lancet (1972) 1:1393. doi: 10.1016/s0140-67367291123-3
31. Gerner RH, Fairbanks L, Anderson GM, Young JG, Scheinin M, Linnoila M, et al. CSF Neurochemistry in Depressed, Manic, and Schizophrenic Patients Compared With That of Normal Controls. Am J Psychiatry (1984) 141:1533–40. doi: 10.1176/ajp.141.12.1533
32. Miller CL, Llenos IC, Cwik M, Walkup J, Weis S. Alterations in Kynurenine Precursor and Product Levels in Schizophrenia and Bipolar Disorder. Neurochem Int (2008) 52:1297–303. doi: 10.1016/j.neuint.2008.01.013
33. Chiaroni P, Azorin JM, Bovier P, Widmer J, Jeanningros R, Barré A, et al. A Multivariate Analysis of Red Blood Cell Membrane Transports and Plasma Levels of L-Tyrosine and L-Tryptophan in Depressed Patients Before Treatment and After Clinical Improvement. Neuropsychobiology (1990) 23:1–7. doi: 10.1159/000118707
34. Hoekstra R, Fekkes D, Loonen AJ, Pepplinkhuizen L, Tuinier S, Verhoeven WM. Bipolar Mania and Plasma Amino Acids: Increased Levels of Glycine. Eur Neuropsychopharmacol (2006) 16:71–7. doi: 10.1016/j.euroneuro.2005.06.003
35. Liu H, Ding L, Zhang H, Mellor D, Wu H, Zhao D, et al. The Metabolic Factor Kynurenic Acid of Kynurenine Pathway Predicts Major Depressive Disorder. Front Psychiatry (2018) 9:552. doi: 10.3389/fpsyt.2018.00552
36. Moller SE, Amdisen A. Plasma Neutral Amino Acids in Mania and Depression: Variation During Acute and Prolonged Treatment With LTryptophan. Biol Psychiatry (1979) 14:131–9.
37. Murata S, Murphy M, Hoppensteadt D, Fareed J, Welborn A, Halaris A. Effects of Adjunctive Inflammatory Modulation on IL-1b in Treatment Resistant Bipolar Depression. Brain Behav Immun (2020) 87:369–76. doi: 10.1016/j.bbi.2020.01.004
38. Pan JX, Xia JJ, Deng FL, Liang WW, Wu J, Yin BM, et al. Diagnosis of Major Depressive Disorder Based on Changes in Multiple Plasma Neurotransmitters: A Targeted Metabolomics Study. Transl Psychiatry (2018) 8:130. doi: 10.1038/s41398-018-0183-x
39. Platzer M, Dalkner N, Fellendorf FT, Birner A, Bengesser SA, Queissner R, et al. Tryptophan Breakdown and Cognition in Bipolar Disorder. Psychoneuroendocrinology (2017) 81:144–50. doi: 10.1016/j.psyneuen. 2017.04.015
40. Poletti S, Melloni E, Aggio V, Colombo C, Valtorta F, Benedetti F, et al. Grey and White Matter Structure Associates With the Activation of the Tryptophan to Kynurenine Pathway in Bipolar Disorder. J Affect Disord (2019) 259:404–12. doi: 10.1016/j.jad.2019.08.034
41. Savitz J, Dantzer R, Wurfel BE, Victor TA, Ford BN, Bodurka J, et al. Neuroprotective Kynurenine Metabolite Indices are Abnormally Reduced and Positively Associated With Hippocampal and Amygdalar Volume in Bipolar Disorder. Psychoneuroendocrinology (2015) 52:200–11. doi: 10.1016/ j.psyneuen.2014.11.015

42. Zhou Y, Zheng W, Liu W, Wang C, Zhan Y, Li H, et al. Antidepressant Effect of Repeated Ketamine Administration on Kynurenine Pathway Metabolites in Patients With Unipolar and Bipolar Depression. Brain Behav Immun (2018) 74:205–12. doi: 10.1016/j.bbi.2018.09.007
43. Anderson G, Maes M. Metabolic Syndrome, Alzheimer Disease, Schizophrenia, and Depression: Role for Leptin, Melatonin, Kynurenine Pathways, and Neuropeptides. In: Faroqui A, Faroqui T, editors. M., Syndome and Neurolgical Disorders. Wiley (2013). p. 235–248.
44. Garrison AM, Parrott JM, Tuñon A, Delgado J, Redus L, O'Connor JC. Kynurenine Pathway Metabolic Balance Influences Microglia Activity: Targeting Kynurenine Monooxygenase to Dampen Neuroinflammation. Psychoneuroendocrinology (2018) 94:1–10. doi: 10.1016/j.psyneuen.2018.04.019
45. De Picker LJ, Morrens M, Chance SA, Boche D. Microglia and Brain Plasticity in Acute Psychosis and Schizophrenia Illness Course: A Meta-Review. Front Psychiatry (2017) 16(8):238. doi: 10.3389/fpsyt.2017.00238
46. Miller BJ, Buckley P, Seabolt W, Mellor A, Kirkpatrick B. Meta-Analysis of Cytokine Alterations in Schizophrenia: Clinical Status and Antipsychotic Effects. Biol Psychiatry (2011) 70:663–71. doi: 10.1016/j.biopsych. 2011.04.013
47. De Picker L, Fransen E, Coppens V, Timmers M, de Boer P, Oberacher H, et al. Immune and Neuroendocrine Trait and State Markers in Psychotic Illness: Decreased Kynurenines Marking Psychotic Exacerbations. Front Immunol (2020) 10:2971. doi: 10.3389/fimmu.2019.02971
48. Fukui S, Schwarcz R, Rapoport SI, Takada Y, Smith QR. Blood–Brain Barrier Transport of Kynurenines: Implications for Brain Synthesis and Metabolism. J Neurochem (1991) 56:2007–17. doi: 10.1111/j.1471-4159.1991.tb03460.x
49. Bay-Richter C, Linderholm KR, Lim CK, Samuelsson M, Träskman-Bendz L, Guillemin GJ, et al. A Role for Inflammatory Metabolites as Modulators of the Glutamate N-Methyl-D-Aspartate Receptor in Depression and Suicidality. Brain
50. Raison CL, Dantzer R, Kelley KW, Lawson MA,Woolwine BJ, Vogt G, et al. CSF Concentrations of Brain Tryptophan and Kynurenines During Immune Stimulation With IFN-a: Relationship to CNS Immune Responses and Depression. Mol Psychiatry (2010) 15:393–403. doi: 10.1038/mp.2009.116
51. Belladonna ML, Grohmann U, Guidetti P, Volpi C, Bianchi R, Fioretti MC, et al. Kynurenine Pathway Enzymes in Dendritic Cells Initiate Tolerogenesis in the Absence of Functional IDO. J Immunol (2006) 177:130–7. doi: 10.4049/jimmunol.177.1.130
52. Huang L, Baban B, Johnson BA, Mellor AL. Dendritic Cells, Indoleamine 2,3 Dioxygenase and Acquired Immune Privilege. Int Rev Immunol (2010) 29 (2):133–55. doi: 10.3109/08830180903349669
53. Salter M, Pogson CI. The Role of Tryptophan 2,3-Dioxygenase in the Hormonal Control of Tryptophan Metabolism in Isolated Rat Liver Cells: Effects of Glucocorticoids and Experimental Diabetes. Biochem J (1985) 229:499–504. doi: 10.1042/bj2290499

54. Askland K, Parsons M. Toward a Biaxial Model of "Bipolar" Affective Disorders: Spectrum Phenotypes as the Products of Neuroelectrical and Neurochemical Alterations. J Affect Disord (2006) 94(1–3):15–33. doi: 10.1016/j.jad.2006.02.024

CHAPTER 4

The role of kynurenines in cognitive dysfunction in bipolar disorder

Hebbrecht K, Morrens M, Giltay EJ, van Nuijs ALN, Sabbe B, van den Ameele S.

Neuropsychobiology. 2022;81(3):184-191.

Abstract

Introduction: Chronic low-grade inflammation is suggested to play a pathophysiological role in bipolar disorder (BD) and its related cognitive dysfunctions. Although kynurenine (KYN) pathway metabolites are key inflammatory mediators, studies investigating the association between KYN metabolism and cognition in BD are scarce. We aimed to explore the relationship between KYN metabolism and cognitive functioning across different mood states in BD.

Methods: Sixty-seven patients with BD (35 depressed and 32 [hypo] manic) and 29 healthy controls were included. Cognitive functioning was assessed at 3 time intervals (baseline, 4, and 8 months) assessing processing speed, sustained attention, verbal memory, working memory, and response inhibition. Plasma samples for quantification of 3-hydroxykynurenine, quinolinic acid, and kynurenic acid (KYNA) were concurrently provided. Linear mixed models were used for statistical analysis.

Results: The manic group showed deficits in all assessed cognitive domains with the exception of verbal memory at all test moments. The bipolar depression group showed deficits in the processing speed at all test moments. Throughout the whole follow-up period, KYNA was significantly lower in both patient groups than in controls. Only in the bipolar depression group, low KYNA was associated with worse global cognitive functioning (B = 0.114, $p = 0.02$) and slower processing speed in particular (B = 0.139, $p = 0.03$).

Conclusion: Only in the bipolar depression group, lower KYNA was associated with worse cognitive functioning. Future large-scale longitudinal studies are warranted to confirm the role of KYN metabolites in cognitive impairment in patients with BD and the possible therapeutic implications of this relationship.

Introduction

Apart from mood symptoms, bipolar disorder (BD) is often characterized by highly disabling cognitive impairments that may persist in euthymic phases [1–3]. In particular, domains like verbal memory, executive functioning, processing speed, and attention have been shown to be impaired [4–6], but the specific profile and severity of impairments are heterogeneous between patients [7, 8]. The cognitive decline seems to be associated with the duration and severity of clinical illness and hereby supports the idea of neuroprogression in BD [9–11]. This theoretical construct proposes that, due to subsequent mood episodes, adaptive mechanisms of the body would get exhausted, which may result in maladaptive biological processes, accumulation of damage, and increasing clinical and cognitive burden [12, 13]. A recently published meta-analysis however has questioned the neuroprogression hypothesis in favor of a neurodevelopmental etiology by indicating a stable longitudinal course of cognitive functioning in BD patients [14]. Overall, a better understanding of the multifactorial etiopathogenesis of cognitive impairments in BD is needed. Unraveling the biological processes underlying neurocognitive deficits in BD will play a key role in this understanding.

Recent evidence shows that the kynurenine (KYN) pathway, upregulated by inflammation, giving rise to several neuroactive metabolites (online suppl. Fig. 1; see www.karger.com/doi/10.1159/000520152, for all online suppl. material), is dysregulated in BD [15]. Alterations in KYN metabolism have been put forward as a possible cause of cognitive deficits [16]. Disbalances between putatively neurotoxic and neuroprotective KYN metabolites leading to brain damage and cognitive deficits have been described in rodents [17, 18]. Furthermore, an increased metabolism down the neurotoxic branch of the KYN pathway has also been suggested in neurodegenerative disorders such as Alzheimer's disease [16]. In healthy elderly, activation of the KYN pathway was associated with poorer cognitive performance [19]. Studies on cognition and KYN metabolism in psychiatric populations are scarce, but some initial studies in MDD [20, 21] and BD [22] showed some evidence for an association between elevated levels of supposedly neurotoxic metabolites (3 hydroxykynurenine [3-HK], quinolinic acid [QA]), and lower levels of neuroprotective metabolites (kynurenic acid [KYNA]), and cognitive impairment. No previous study has investigated the association between KYN metabolites and cognitive functioning throughout different mood states in BD. However, 2 recently published meta-analyses have shown distinct KYN metabolite alterations between different mood states in BD. Marx et al. [23] revealed in their meta-analysis distinct alterations between bipolar and manic patients with a decrease in KYN and QA only occurring in the depressed group and a decrease in tryptophan and KYNA only occurring in acute states (i.e., mania or depression) but not in euthymia. Bartoli et al. [15] found similar KYN metabolite alterations in depressed and manic BD patients, but the reduction in KYNA levels was more pronounced in the depressed group, and the

reduction in tryptophan levels more pronounced in the manic group. Above-described results of subgroup analyses were however based on a limited number of studies.

The aim of this study was to investigate the impact of KYN metabolite levels on cognitive performance and their relation to mood symptoms in depressed and (hypo)manic patients with BD. We hypothesized that an increased level of the putatively neurotoxic (3-HK and QA) and a decreased level of the neuroprotective (KYNA) end products would be more pronounced during acute mood episodes and related to increased cognitive deficits. By elucidating this mood state-specific relation between KYN metabolism and cognitive functioning, we aimed to gain insight into the pathophysiology of cognitive deficits in BD and the possible mediating role of KYN metabolites therein. In the long term, more insight in these processes may offer new opportunities regarding treatment and monitoring of illness course.

Material and Methods

Participants

This study was performed in the context of a longitudinal research project named "Inflammation and Neuroprotection in Bipolar Disorder" [24]. The study sample consisted of adults of 18–65 years of age with a DSM-IV-TR diagnosis of BD type I, type II, or schizoaffective disorder and gender- and age-matched healthy controls (HCs). Patients or controls with following criteria were excluded: substance abuse, use of anti-inflammatory drugs within 2 weeks before screening, acute infection, autoimmune diseases, chronic inflammatory or neurological diseases, pregnancy or breastfeeding, electroconvulsive therapy within 6 months before screening or during follow-up, mental retardation, disturbances on screening blood test (blood count; electrolytes; fasting glucose; lipid profile; and liver, kidney, and thyroid function ad serology). Additional exclusion criteria for the control group were current or past diagnosis of DSM-IV-TR MDD, BD or psychotic syndrome, BD or psychotic syndrome in first-degree relatives, and current use of psychopharmacological drugs. The study was approved by the Committee for Medical Ethics of the University Hospital Antwerp and the Antwerp University with protocol number B300201421645. The local Ethical Committees of the participating centers approved the protocol. All participants agreed to participate in the study and signed informed consent. The study complied with the Declaration of Helsinki.

Study Design

Patients were recruited during an acute mood episode and assigned to either the depressed (BD-D) or (hypo)manic (BD-M) subgroup. In both patients and controls, screening was followed by a first test day after 1–5 days. Two subsequent test days were

planned after 4 and 8 months. Every test day included the same clinical, cognitive, and laboratory assessments as described below. During the study period, patients received treatment as usual without therapeutic intervention of the investigators.

Clinical Assessments

The M.I.N.I.-Plus, International Neuropsychiatric Interview, version 5.0.0 was used to establish diagnoses in patients and rule out axes I diagnoses in HCs [25]. In patients, the severity of mood symptoms was assessed by the 17-item Hamilton Depression Rating Scaling (HDRS-17) [26] and the Young Mania Rating Scale (YMRS) [27]. The threshold score for inclusion was ≥17 for the HDRS-17 [28] or ≥13 for the YMRS [29], corresponding to moderate depression or hypomania, respectively. Psychotic symptoms were evaluated using the positive subscale of the Positive and Negative Syndrome Scale (PANSS) [30]. All clinical assessments were done by a psychiatrist in training (SvdA) and supervised by a psychiatrist (MM).

Cognitive Tasks

Cognitive functioning was assessed by means of 5 subtasks of the International Society of Bipolar Disorders Battery for the Assessment of Neurocognition (ISBD-BANC): Brief Assessment of Cognition in Schizophrenia-Symbol Coding (Y-BACS) to assess the speed of processing; the Continuous Performance Task-Identical Pairs (CPT-IP) to assess vigilance/sustained attention; the Hopkins Verbal Learning Test – Revised (HVLT-R) to evaluate verbal learning and memory; the Letter Number Span (LNS) for working memory; and the Color-Word Interference Test (D-KEFS – CWIT) to assess executive functioning (response inhibition).

Bioanalytical Measurements

Blood was drawn by venipuncture between 08:00 A.M. and 10:30 A.M. Citrate plasma tubes were immediately stored at 4°C, centrifuged at 2,000 g and 4°C for 10 min within 2 h after blood draw, and the plasma was aliquoted and stored at −70°C until assayed. Patient and control samples were analyzed in a randomized sequence. Calibration curves (3-HK, 8–4, 500 nM; QA, 12–6,000 nM; KYNA, 10–5, 500 nM) and sample extracts were analyzed using ultrahigh-pressure liquid-chromatography coupled to tandem mass spectrometry (Agilent 1290 Infinity LC system coupled to an Agilent 6460 Triple Quadrupole MS system). A more extensive elaboration on the sample preparation and UPLC analysis can be found in our previously published articles [24, 31].

Statistical Analysis

Normality of outcome variables and homoscedasticity of residuals were assessed by visual inspection. All KYN parameters were log-transformed to obtain a normal distribution.

Baseline demographic and metabolic characteristics were compared between the BD-D, BD-M, and HC groups by χ^2 tests for categorical variables and one-way ANOVA for continuous variables. Tukey's HSD corrections for multiple comparisons were performed to conduct pairwise comparisons between all considered groups. Differences in baseline cognitive task scores and KYN metabolite levels between the 3 groups were studied using one-way ANOVA with post hoc analysis with Tukey's HSD correction. An additional Bonferroni-Holm correction to the p values was applied in order to correct for multiple testing.

The longitudinal evolution of cognitive task scores and KYN metabolite levels were plotted using linear mixed model (LMM) analysis with the different KYN metabolites and cognitive task scores as the dependent variable. The subject ID was included as random effect, and group, moment and the interaction group x moment were added stepwise as fixed effects. Nonsignificant interaction terms were eliminated from the final model.

Considering the relatively stable impairments in cognitive scores and KYN levels over the 3 test moments, the association between cognition and KYNs was investigated by combining the data of the 3 test moments. Associations between the cognitive task scores and KYN metabolite levels were studied using an LMM that included all 5 standardized cognitive tasks as the dependent variable, subject ID as the random effect, and the group (BD-D, BD-M, and HC) as the fixed effect. In a second step, the different KYN metabolites were also separately added as the fixed effect to the model. Subsequently, similar univariate analyses with the separate cognitive tasks as the dependent variable were performed. All LMM analyses assessing the relationship between KYN and cognition were additionally adjusted for sex, age, smoking, body mass index, and years of education (by adding these as covariates to the model). All statistical analyses were performed in JMP Pro 14 (JMP, Marlow, UK). R statistical software was used for the graphical illustrations (R version 3.6.0; R Foundation for Statistical Computing, Vienna, Austria, 2016. URL: https://www.project.org/).

Results

Participants

At screening, 35 patients were included in the BD-D group and 32 patients in the BD-M group. Thirty-five HCs were included and matched by age and sex. An overview of demographic and metabolic characteristics of the 3 groups (HC, BD-D, and BD-M) can be found in Table 1.

Table 1. Demographic and metabolic characteristics of participants

	HC (n=35)	**BD-D** (n=35)	**BD-M** (n=32)	Statistics†
Gender (n (%))	19 (54.3)	24 (68.6)	15 (46.9)	n.s.
Age, years (mean ± SD)	42.7 ± 11.6	43.7 ± 9.7	42.4 ± 12.7	n.s.
Caucasian (n (%))	34 (97.1)	33 (94.3)	30 (93.8)	n.s.
Smokers (n (%))	6 (17.1)	18 (51.4)	14 (43.8)	0.006
BMI (mean ± SD), kg/cm²	23.7 ± 2.6	24.3 ± 3.5	26.3 ± 4.7	0.013‡
Waist(mean ± SD), cm	84.4 ± 9.6	87.1 ± 11.7	91.2 ± 12.4	n.s.
Fasting glucose (mean ± SD), mg/dl	87.5 ± 6.8	88.9 ± 8.8	91.1 ± 9.8	n.s.
Cholesterol (mean ± SD), mg/dl				
Total	190.9 ± 42.4	189.1 ± 42.7	183.7 ± 47.3	n.s.
HDL	62.0 ± 18.4	61 ± 15.6	56.0 ± 19.9	n.s.
LDL	109.0 ± 33.4	108.4 ± 41.5	102.4 ± 40.6	n.s.

Data presented as mean ± SD (range) or n (%).
HC, healthy controls; BD-D, patients with bipolar disorder and depression at baseline; BD-M, patients with bipolar disorder and (hypo)mania at baseline; BMI, body mass index; HDL, high density lipoprotein; LDL, low density lipoprotein.
†P-values of one-way ANOVA for continuous data and chi-square for categorical. P-values < .05 are italicized; n.s. = not significant (P> .05).
‡Tukey HSD: BMI: BD-M > HC; Albumin: BD-M < HC

Clinical characteristics of the 2 patient groups are summarized in Table 2. Twenty-nine BD-D patients, 20 BD-M patients, and 30 controls completed the 4- and 8-month follow-up. More information on dropout can be found in our previously published article [24].

Table 2. Clinical characteristics and baseline data of patients

	BD-D (n=35)	**BD-M** (n=32)
Diagnosis		
BD type 1	16 (45.7)	26 (81.3)
BD type 2	19 (54.3)	4 (12.5)
Schizoaffective disorder	0	2 (6.3)
Age of onset, years	22.5 (10.9)	27.5 (11.8)
Duration of illness, years	20.6 (10.2)	14.3 (11.8)
First episode: depression	23 (67.7)	17 (53.1)
Lifetime psychotic features	13 (37.1)	24 (75.0)
Total number of hospitalizations		
0	6 (17.1)	3 (9.4)
1-5	23 (65.7)	22 (68.8)
6-10	6 (17.1)	7 (21.9)
Lifetime substance abuse		
Alcohol	11 (31.4)	7 (21.8)
THC	6 (17.1)	7 (21.8)
Hard drugs	4 (11.4)	1 (3.1)
Baseline medication use		
Medication-free	3 (8.6)	3 (9.4)
Lithium	13 (37.1)	11 (34.4)
Valproate	6 (17.1)	3 (9.4)
Carbamazepine	1 (2.9)	2 (6.3)
Lamotrigine	7 (20.0)	1 (3.1)
Antipsychotic	19 (54.3)	23 (71.9)
Antidepressant	25 (71.4)	6 (18.8)
Benzodiazepine	8 (22.9)	16 (50.0)
Baseline mood scales		
HDRS-17	21.2 (4.0)	8.1 (5.7)
YMRS	5.5 (4.8)	20.9 (5.1)
PANSS positive subscale	9.8 (3.4)	15.0 (4.7)
Psychotic features present	5 (14.3)	12 (37.5)

Data presented as mean ± SD (range) or n (%).
BD-D, patients with bipolar disorder and depression at baseline; BD-M, patients with bipolar disorder and (hypo)mania at baseline; BD, bipolar disorder; THC, tetrahydrocannabinol; HDRS-17, 17-item Hamilton depression rating scale; YMRS, Young mania rating scale; PANSS, Positive and Negative Syndrome Scale.

Evolution of Mood Symptoms

The BD-M group showed a rapid decline of manic symptoms with a YMRS of 5.4 (SD 6.3) at 4 months and 5.5 (SD 6.39) at 8 months. The BD-D group showed a chronic subsyndromal depressive level with an HDRS score at 4 months of 13.4 (SD 6.2) and at 8 months of 12.4 (SD 7.4).

Evolution of Cognitive and KYN Measurements

Baseline differences in cognitive and KYN measurements for the 3 groups are described in online supplementary material. The BD-M group scored significantly lower on the BACS-SC ($F[2] = 11.3$; $p < 0.001$), LNS ($F[2] = 4.5$; $p = 0.042$), CPT-IP ($F[2] = 8.8$; $p = 0.001$), and CWT-IT ($F[2] = 4.5$; $p = 0.042$) than the HC group at baseline (online suppl. Table 1). These scores remained significantly lower than HC scores throughout the whole follow-up, except for the CWT-IT which gradually improved over time (reaching to scores comparable to the BD-D and HC group) (online suppl. Fig. 2). In the BD-D group, only a significant impairment in the BACS-SC score compared to HC at baseline ($F[2] = 11.3$; $p < 0.001$) was seen (online suppl. Table 1), and this difference remained also stable over time (BD-M, BD-D < HC; $F[2] = 11.3$; $p < 0.0001$) (online suppl. Fig. 2). There were no significant differences in the HVLT-R score between the BD-D, BD-M, and HC groups at baseline ($F[2] = 1.6$; $p = 0.207$) (online suppl. Table 1). The 3 groups however showed a different evolution over time (significant group x time interaction $F[4] = 3.0$; $p = 0.021$) (online suppl. Fig. 2). Baseline KYNA concentration was significantly lower in the BD-D and BD-M groups than the HC group ($F[2] = 5.5$; $p = 0.022$) (online suppl. Table 2). These between group differences in KYNA remained significantly lower over time in the BD-M and BD-D groups than the HC group ($F[2] = 8.0$; $p = 0.0006$) (online suppl. Fig. 3).

Association between Cognition and Markers of KYN Metabolism

As shown in Figure 1, a positive association between KYNA levels and the score on the overall cognitive test battery, using data from all test moments, was found in the BD-D group ($B = 0.114$, $p = 0.02$). This positive association was mainly true for BACS-SC ($B = 0.139$, $p = 0.03$). In the BD-M and HC groups, no significant associations were found between the overall cognitive test battery and any of the KYN metabolites (online suppl. Fig. 4; all $p > 0.05$ for BD-M and HC groups). In the BD-M group, KYNA was also positively associated with an improved BACS-SC score ($B = 0.198$, $p = 0.006$) ($B = 0.279$, $p = 0.02$).

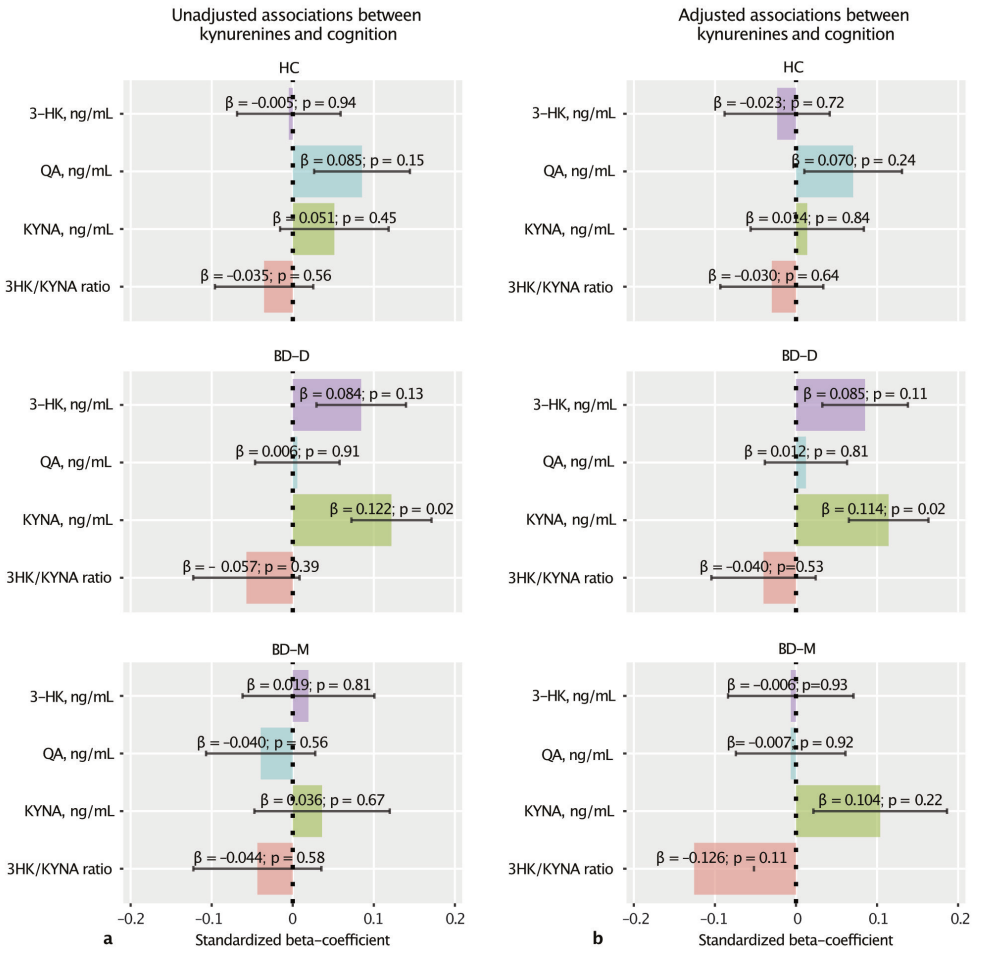


Figure 1. Association between the global cognitive task score and KYN metabolites, as analyzed by using LMMs.

a Overview of associations between the cognitive task scores and KYN metabolite levels as analyzed by LMMs that included all 5 standardized cognitive tasks as the dependent variable, the subject ID as the random effect, and the group (BD-D, BD-M, and HC) as the fixed effect. **b** LMMs adjusted for sex, age, smoking, BMI, and years of education. BD-D, patients with bipolar disorder and depression at baseline; BD-M, patients with bipolar disorder and (hypo)mania at baseline; BD, bipolar disorder; 3-HK, 3-hydroxykynurenine; QA, quinolinic acid; KYNA, kynurenic acid; LMM, linear mixed model; BMI, body mass index; KYN, kynurenine.

Discussion

This study investigated the association between cognitive functioning and KYN metabolites in patients with BD in different mood states. Comparable to the previous research [4–6], we found a broad range of cognitive deficits in patients with BD compared to HCs. BD-M patients showed impairments in processing speed, sustained attention, inhibition, and working memory, and these remained impaired after remission from the manic episode, except for the improvement of response inhibition after remission. BD-D patients showed deficits in the processing speed and also remained impaired in the chronic mild depressed states that characterized this group at the follow-up. Stable low levels of KYNA were found in BD patients, independent of their mood state. The most consistent finding supporting our hypothesis was the association between lower KYNA levels and lower scores on the overall cognitive test battery but only in the BD-D group.

KYNA is generally considered to have neuroprotective effects, but the exact role of KYNA in both healthy and pathological conditions remains poorly understood [16]. KYNA is an antagonist of the N-methyl-D-aspartate receptor and the alpha7 nicotinic receptor, leading to alterations in glutamatergic and cholinergic signaling [16,32]. Decreased KYNA levels in BD have been described several times and confirmed by a recent meta-analysis [15]. In the only previous study investigating the relationship between KYN metabolites and cognition, a higher ratio of 3 HK/KYNA in euthymic men with BD was related to impaired verbal learning and memory, explained by the decreased neuroprotective impact of KYNA [22]. The stable low KYNA level in patients and the positive correlation between KYNA and cognitive scores in patients with depressive symptoms in our study could suggest a critical low concentration from which KYNA's neuroprotective effects are insufficient to protect for neuronal damage and consequently cause decreased cognitive functioning. Previous studies in MDD have also found structural evidence for a relationship between KYNA and cognition as they found a negative correlation between the KYNA/3-HK ratio and hippocampal activity and a positive correlation of KYNA/QA with memory recall and hippocampal and amygdala volumes [21, 33]. It is probable that the KYN pathway influences glutamatergic neurotransmission in brain regions that are in association with cognitive tasks [34].

We however want to emphasize the hypothetical character of this interpretation since several findings do not support the relation between cognitive impairment and decreased KYNA levels. First, no association between cognition and KYN metabolism could be found in the, more cognitively impaired, BD-M group. Second, in the BD-D group, not all cognitive tests were associated with a lower KYNA level. This was mainly true for the processing speed (as assessed by BACS-SC). Again, even if KYN metabolites are a weak proxy of cognitive functioning in BD, other factors seem to be more determining.

Interestingly, we did not observe a clear state pattern of KYN metabolite levels. In particular, no clear difference in KYN metabolite levels was observed between the acute manic state and after remission. This finding was in contrast to a previous study of our group in patients with psychotic disorder in which a state-specific decrease of KYNA was found during the acute psychotic state [35]. As for the BD-D group, we should refrain from any interpretation on state/trait markers considering the chronic mild depressed state in the follow-up period. The time frame between the acute mood episode, the accompanying biological alterations, the possible neurotoxic effects, and the subsequent translation into a change in cognitive test scores is unknown. To formally exclude a potential role of KYN metabolism in cognitive deterioration, the impact of just 1 mood episode with a follow-up period of 8 months may possibly be too limited.

Our study has many strengths. The repeated measurements within the same study sample increased the power of our analysis and provided an internal control. The impact of methodological bias was minimized by standardized blood sampling, randomization of samples, and uniform, meticulous laboratory procedures. All cognitive and clinical assessments were done by the same clinician-researcher, excluding inter-rater bias. We provided answers on an a priori defined research question without "cherry-picking" significant results in order to represent a clear and realistic view of the evidence, and we used rigorous statistical methods with correction for multiple testing.

There are some noteworthy limitations to our study. First, although 3-HK easily crosses the blood-brain barrier, the passage of QA and KYNA is more complex [16]. Findings in peripheral samples possibly do not represent a direct quantitative representation of central processes. However, peripheral measurements may be of great value in view of monitoring disease activity and staging. Second, although we adjusted for some important confounding variables known to influence KYN metabolites and cognitive functioning (age, sex, body mass index, smoking, and years of education), it was not possible to adjust for all potential confounders (such as severity of mood symptoms and pharmacological treatments). Third, combining the data on all 3 test moments precluded us to make statements on the temporal relation between KYN level alterations and cognitive impairment. Furthermore, as mood normalized for a subgroup of patients throughout the follow-up period (more in the manic group than in the chronic depressed group), it is possible that the strengths of the associations between KYN metabolite alterations and mood states were somewhat attenuated due to combining the data of different test moments. Fourth, as cognitive impairments are mainly considered as a trait effect, cognitive measurements during affective states could be influenced by mood symptoms [36]. To completely rule out the impact of mood symptoms on cognitive functioning, an investigation of KYN metabolites and cognition in euthymic patients would be useful.

In conclusion, KYNA, with theoretically neuroprotective properties, showed low levels in patients independent of their mood state. A mild positive association was found between KYNA levels and the overall cognitive test battery but only in the BD-D group. In view of the heterogeneity regarding the patient profile, course of illness and cognitive deficits in BD, long-term studies with detailed observation of the mood state, cognitive function, and biological correlates are required to further illuminate the state-related dynamics of KYN metabolites and their relationship with cognition over time.

Supplementary Material

See www.karger.com/doi/10.1159/000520152, for all online supplementary material

References

1. Brissos S, Dias VV, Kapczinski F. Cognitive performance and quality of life in bipolar disorder. Can J Psychiatry. 2008; 53(8): 517–24.
2. Martinez-Aran, Vieta E, Reinares M, Colom F, Torrent C, S.nchez-Moreno J, et al. Cognitive function across manic or hypomanic, depressed, and euthymic states in bipolar disorder. Am J Psychiatry. 2004 Feb; 161(2): 262–70.
3. Ceylan D, Akdede BB, Bora E, Aktener AY, Hıdıroğlu Ongun C, Tunca Z, et al. Neurocognitive functioning during symptomatic states and remission in bipolar disorder and schizophrenia: a comparative study. Psychiatry Res. 2020 Oct 1; 292: 113292.
4. Bora E, Yücel M, Pantelis C, Berk M. Metaanalytic review of neurocognition in bipolar II disorder. Acta Psychiatr Scand. 2011; 123: 165–74.
5. Arts B, Jabben N, Krabbendam L, van Os J. Meta-analyses of cognitive functioning in euthymic bipolar patients and their first-degree relatives. Psychol Med. 2008 Jun 9; 38(6): 771–85.
6. Kurtz MM, Gerraty RT. A meta-analytic investigation of neurocognitive deficits in bipolar illness: profile and effects of clinical state. Neuropsychology. 2009; 23(5): 551–62.
7. Bora E. Differences in cognitive impairment between schizophrenia and bipolar disorder: Considering the role of heterogeneity. J Affect Disord. 2018.
8. Van Rheenen T, Lewandowski K, Tan E, Ospina L, Ongur D, Neill E. Characterizing cognitive heterogeneity on the schizophrenia-bipolar disorder spectrum. Psychol Med. 2017; 47(10): 1848–64.
9. Bora E. Neurocognitive features in clinical subgroups of bipolar disorder: a meta-analysis. J Affect Disord. 2018; 229: 125–34. Elsevier B.V.
10. Cardoso T, Bauer IE, Meyer TD, Kapczinski F, Soares JC. Neuroprogression and cognitive functioning in bipolar disorder: a systematic review. Curr Psychiatry Rep. 2015; 17: 75. Current Medicine Group LLC 1.
11 Lewandowski KE, Cohen BM, .ngur D. Evolution of neuropsychological dysfunction during the course of schizophrenia and bipolar disorder. Psychol Med. 2011 Feb; 41(2): 225 41.
12. Berk M, Kapczinski F, Andreazza AC, Dean OM, Giorlando F, Maes M, et al. Pathways underlying neuroprogression in bipolar disorder: focus on inflammation, oxidative stress and neurotrophic factors. Neurosci Biobehav Rev. 2011; 35: 804–17.
13. Berk M, Post R, Ratheesh A, Gliddon E, Singh A, Vieta E, et al. Staging in bipolar disorder: from theoretical framework to clinical utility. World Psychiatry. 2017 Oct 1; 16(3): 236–44.
14. Szmulewicz A, Valerio MP, Martino DJ. Longitudinal analysis of cognitive performances in recent-onset and late-life bipolar disorder: a systematic review and meta-analysis. Bipolar Disord. 2020 Feb 2; 22(1): 28–37.
15. Bartoli F, Misiak B, Callovini T, Cavaleri D, Cioni RM, Crocamo C, et al. The kynurenine pathway in bipolar disorder: a meta-analysis on the peripheral blood levels of tryptophan and related metabolites. Mol Psychiatry. 2020; 26(7): 3419–29.
16. Savitz J. The kynurenine pathway: a finger in every pie. Mol Psychiatry. 2020; 25: 131 47.

17. Alexander KS, Wu HQ, Schwarcz R, Bruno JP. Acute elevations of brain kynurenic acid impair cognitive flexibility: normalization by the alpha7 positive modulator galantamine. Psychopharmacology. 2012 Apr; 220(3): 627–37.
18. Erhardt S, Schwieler L, Emanuelsson C, Geyer M. Endogenous kynurenic acid disrupts prepulse inhibition. Biol Psychiatry. 2004 Aug 15; 56(4): 255–60.
19. Solvang SH, Nordrehaug JE, Tell GS, Nygard O, McCann A, Ueland PM, et al. The kynurenine pathway and cognitive performance in community-dwelling older adults. The Hordaland Health Study. Brain Behav Immun. 2019 Jan 1; 75: 155–62.
20. Zhou Y, Zheng W, Liu W, Wang C, Zhan Y, Li H, et al. Cross-sectional relationship between kynurenine pathway metabolites and cognitive function in major depressive disorder. Psychoneuroendocrinology. 2019 Mar 1; 101: 72–9.
21. Young KD, Drevets WC, Dantzer R, Teague TK, Bodurka J, Savitz J. Kynurenine pathway metabolites are associated with hippocampal activity during autobiographical memory recall in patients with depression. Brain Behav Immun. 2016 Aug 1; 56: 335–42.
22. Platzer M, Dalkner N, Fellendorf FT, Birner A, Bengesser SA, Queissner R, et al. Tryptophan breakdown and cognition in bipolar disorder. Psychoneuroendocrinology. 2017 Jul; 81: 144–50.
23. Marx W, McGuinness AJ, Rocks T, Ruusunen A, Cleminson J, Walker AJ, et al. The kynurenine pathway in major depressive disorder, bipolar disorder, and schizophrenia: a meta-analysis of 101 studies. Mol Psychiatry. 2020.
24. van den Ameele S, van Nuijs AL, Lai FY, Schuermans J, Verkerk R, van Diermen L, et al. A mood state-specific interaction between kynurenine metabolism and inflammation is present in bipolar disorder. Bipolar Disord. 2019 Feb; 22(1): 59–69.
25. Lecrubier Y, Sheehan D, Weiller E, Amorim P, Bonora I, Sheehan KH, et al. The mini international neuropsychiatric interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. Eur Psychiatry. 1997 Jan 1; 12(5): 224–31.
26. Trajković G, Starčević V, Latas M, Leštarević M, Ille T, Bukumirić Z, et al. Reliability of the Hamilton rating scale for depression: a metaanalysis over a period of 49 years. Psychiatry Res. 2011 Aug 30; 189(1): 1–9.
27. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. Br J Psychiatry. 1978 Nov; 133: 429–35.
28. Zimmerman M, Martinez JH, Young D, Chelminski I, Dalrymple K. Severity classification on the Hamilton depression rating scale. J Affect Disord. 2013 Sep 5; 150(2): 384–8.
29. Vieta E. Guide to assessment scales in bipolar disorder. Tarporley: Springer Healthcare Ltd; 2010.
30. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull. 1987; 13(2): 261–76.
31. van den Ameele S, Fuchs D, Coppens V, de Boer P, Timmers M, Sabbe B, et al. Markers of inflammation and monoamine metabolism indicate accelerated aging in bipolar disorder. Front Psychiatry. 2018; 9: 250.

32. Erhardt S, Schwieler L, Imbeault S, Engberg G. The kynurenine pathway in schizophrenia and bipolar disorder. Neuropharmacology. 2017; 112: 297–306.
33. Savitz J, Drevets WC, Smith CM, Victor TA, Wurfel BE, Bellgowan PS, et al. Putative neuroprotective and neurotoxic kynurenine pathway metabolites are associated with hippocampal and amygdalar volumes in subjects with major depressive disorder. Neuropsychopharmacology. 2015; 40(2): 463–71.
34. Fourrier C, Singhal G, Baune BT. Neuroinflammation and cognition across psychiatric conditions. CNS Spectr. 2019; 24(1): 4–15.
35. De Picker L, Fransen E, Coppens V, Timmers M, de Boer P, Oberacher H, et al. Immune and neuroendocrine trait and state markers in psychotic illness: decreased kynurenines marking psychotic exacerbations. Front Immunol. 2020 Jan 17; 10: 2971.
36. Miskowiak KW, Burdick KE, Martinez-Aran A, Bonnin CM, Bowie CR, Carvalho AF, et al. Methodological recommendations for cognition trials in bipolar disorder by the International Society for Bipolar Disorders Targeting Cognition Task Force. Bipolar Disord. 2017 Dec 1; 19(8): 614–26.

CHAPTER 5

Understanding personalized dynamics to inform precision medicine: a dynamic time warp analysis of 255 depressed inpatients

Hebbrecht K, Stuivenga M, Birkenhäger T, Morrens M, Fried EI, Sabbe B, Giltay EJ.

BMC Med. 2020 Dec 23;18(1):400.

Abstract

Background: Major depressive disorder (MDD) shows large heterogeneity of symptoms between patients, but within patients, particular symptom clusters may show similar trajectories. While symptom clusters and networks have mostly been studied using cross-sectional designs, temporal dynamics of symptoms within patients may yield information that facilitates personalized medicine. Here, we aim to cluster depressive symptom dynamics through dynamic time warping (DTW) analysis.

Methods: The 17-item Hamilton Rating Scale for Depression (HRSD-17) was administered every 2 weeks for a median of 11 weeks in 255 depressed inpatients. The DTW analysis modeled the temporal dynamics of each pair of individual HRSD-17 items within each patient (i.e., 69,360 calculated "DTW distances"). Subsequently, hierarchical clustering and network models were estimated based on similarities in symptom dynamics both within each patient and at the group level.

Results: The sample had a mean age of 51 (SD 15.4), and 64.7% were female. Clusters and networks based on symptom dynamics markedly differed across patients. At the group level, five dynamic symptom clusters emerged, which differed from a previously published cross-sectional network. Patients who showed treatment response or remission had the shortest average DTW distance, indicating denser networks with more synchronous symptom trajectories.

Conclusions: Symptom dynamics over time can be clustered and visualized using DTW. DTW represents a promising new approach for studying symptom dynamics with the potential to facilitate personalized psychiatric care.

Introduction

Depression is defined by its symptoms (such as a sad mood and insomnia) that are correlated with each other. The dominant explanation in the field has been that these relations stem from a shared causal origin, a perspective termed the common cause framework [1, 2]. The contemporary conceptualization for major depressive disorder (MDD) is similar to that of other medical conditions in that it assumes all observable depressive symptoms are caused by an underlying disease construct [1, 3]. In research, symptoms are usually added up to sum scores, and thresholds are used to indicate case status. This approach assumes that symptoms are equivalent, causally independent, and roughly interchangeable indicators of the underlying disease construct [4]. This conceptual framework has dominated depression research over the past decades: the inclusion criteria in research studies were based on the syndromal DSM diagnoses of MDD, and the unweighted sum scores of depression rating scales were used as a measure for severity and treatment response [5] (e.g., Hamilton Rating Scale for Depression [HRSD] [6] and MontgomeryÅsberg Depression Rating Scale [MADRS]) [7]. However, years of research have shown slow progress in the search for the underlying risk factors and biomarkers of the unitary construct MDD [4]. The need for a new, scientifically sound approach for conceptualizing depression is warranted.

Increasing evidence points towards the multidimensional character of MDD with a high degree of symptomatic variability between and within patients [4, 8]. Individual symptoms are mutually interacting and causing each other [9], and they have different risk factors [10, 11], underlying biology [11–13], psychosocial impact [14], and course trajectories [5]. Recent years have therefore seen a shift in the conceptualization of depression towards a network perspective where the depressive syndrome is hypothesized to stem from mutual causal relations among components of the system, such as depression symptoms [9, 15]. Furthermore, patients manifest specific depression symptom profiles with preferential responses to different treatments. Consequently, there is increasing recognition of the importance of investigating individual symptoms and their timely evolution, both within individual patients and in groups of patients [16, 17]. This is also in line with the aims of the Research Domain Criteria (RDoC) project to deconstruct psychiatric disorders by analyzing the dynamics (e.g., symptom trajectories over time) that lie at their basis [18, 19].

Several factor analytic studies of the HRSD-17 have tried to tackle the symptomatic diversity of MDD by means of identifying homogeneous symptom groups within MDD. Although there was evidence for a "general depression" and "insomnia" factor, the overall results were inconsistent, with reported factors ranging from two to eight [20–23]. Furthermore, factors seemed to change over time [24] and were poorly generalizable to other populations. Hierarchical cluster analysis is another statistical method to decompose MDD into homogeneous symptom groups, and comparable results ("general depression,"

"insomnia") have been found with this approach [25,26]. Network analysis is a more recent approach that expands further on studying the symptom correlations by investigating the influence of symptoms on each other [9]. Both factor and network analyses were mostly conducted on cross-sectional data, and consequently, they did not take the temporal dynamics of symptoms into account. Furthermore, both techniques almost exclusively studied aggregated patient data without studying the intraindividual symptom heterogeneity.

Routine outcome monitoring (ROM) entails the collection of clinical data at baseline and at regular time intervals thereafter in order to monitor disease severity as well as the clinical course during treatment. ROM may provide feedback to both the clinician and the patient and enable "patient-centered research" [27, 28]. Timeseries ROM data enable the capturing of dynamics of symptoms over time using dynamic time warping (DTW). DTW is a widely used statistical algorithm [29, 30], though not yet in psychological and psychiatric research. It is an effective clustering strategy for time-series data across a broad range of application domains [31]. Examples of biomedical applications are speech recognition [16], gait pathology [32–34], and electro-cardiogram analysis [35]. The DTW approach could be well-suited to cluster individual symptoms based on the temporal features that they share, using ROM or ecological momentary assessment EMA [36] time-series data.

In this study, we utilize depression symptom data from a clinical ROM regime, every 2 weeks, of 255 depressed inpatients and present the first implementation of DTW time-series analysis on depression symptom trajectories. This paper is built upon a dual structure in which the DTW analysis is introduced both for intra-individual (i.e., idiographic) and inter-individual (i.e., nomothetic) analysis [37]. In the idiographic analysis, we aim to assess the dynamics and covariation of changes in symptoms over time within each individual patient with two or more assessments and to estimate the symptom clusters and networks within each patient. In the nomothetic analysis, we aim to study the aggregated dynamics of individual symptoms to yield systematic patterns across patients.

Methods

Sample and setting

From the original study sample of 276 consecutive patients (i.e., included in the cohort study in the order that they were admitted), 21 patients had only one HRSD-17 measurement due to a short hospitalization period or refusal to participate, yielding 255 (91.6%) patients included in the current analysis. Thus, we included 255 adult patients consecutively admitted to a tertiary psychiatric hospital in Duffel, Belgium, and fulfilling the MINI-Plus diagnosis, based on the DSM-IV criteria, of a depressive episode as part of

an MDD or bipolar disorder (BD). In order to obtain a representative sample of depressed participants, exclusion criteria were minimal. We did not include patients with (comorbid MINI-Plus) psychotic disorders (including schizoaffective disorder) or with a dependency on alcohol or drugs within 12 months prior to hospitalization. Moreover, patients with insufficient mastery of the Dutch language were not included.

Treatment

Inpatients received treatment as usual which was based on evidence-based guidelines and consisted of pharmacotherapy, (group) psychotherapy, or a combination of both. These guidelines for diagnosis and treatment were formulated by the Dutch Association of Psychiatry, often in association with the associations of psychology and general practitioners (www.trimbos.nl, www.nvvp.net). Psychotropic medication at baseline was coded into five dichotomous variables: antidepressants, mood stabilizers, antipsychotics, benzodiazepines, and stimulants. Response was defined as a ≥ 50% reduction of the HRSD- 17 compared to the baseline assessment. Remission was defined as scoring ≤ 7 on the HRSD-17.

Measurements

The present study was part of a larger follow-up study investigating the feasibility of ROM in the University Psychiatric Centre in Duffel [38]. ROM were done at baseline and every 2 weeks thereafter during the clinical admission which lasted from 2 weeks to 16 months. ROM consisted of a test battery assessing overall mental well-being, quality of life, and mood (including the Hamilton Depression Rating Scale-17). The data presented in this article represent the collected data from the period April 2015 through February 2018.

The HRSD-17 consists of 17 items on a Likert scale, ranging from either 0 to 4 (for 9 items) or 0 to 2 (for 8 items). The internal reliability of the HRSD-17 is adequate with most studies reporting a Cronbach's alpha of ≥ 0.70. It has a good retest and interrater reliability (above 0.80) when assessed over an interval ranging from 1 to 30 days [21]. The Omega and Cronbach's alpha in our sample of 255 patients at baseline were only 0.49 and 0.52, respectively. The Cronbach's alpha improved over time with a score of 0.74 after 2, 0.77 after 4, and 0.79 after 6 weeks. The total score ranges from 0 to 52, and higher scores indicate greater severity, but in the present study, we focus on the trajectories of the 17 individual items only. In order to improve interrater reliability, Hamilton Depression Rating Scale training sessions were organized every 3 months among the in total 6 assessors, during which video-recorded interviews with patients were rated and discussed to reach consensus. In total, they conducted 1480 HRSD-17 assessments in 255 patients, with an average of 5.8 assessments per patient.

Statistical analysis

DTW is an approximate pattern detection algorithm that can measure the similarity between two time-series. It uses a dynamic (i.e., stretching and compressing) programming approach to minimize a predefined distance measure (e.g., Euclidean distance), in order for the two time-series to become optimally aligned through a warping path. The "optimal" alignment minimizes the sum of distances between the aligned elements. The "dtw" (version 1.20.1), "pheatmap" (version 1.0.12), "parallelDist" (version 0.2.4), and "qgraph" (version 1.6.2) packages for the R statistical software were used (R version 3.6.0; R Foundation for Statistical Computing, Vienna, Austria, 2016. URL: https://www.R-project.org/). The idiographic approach per patient was followed by a nomothetic approach to study the depression symptom patterns both within individual patients and in the whole sample of 255 patients. The subsequent methodological steps and statistical methods are described below.

Intra-individual approach

We first aimed to cluster individual symptoms based on the temporal features that they share within each individual patient. The clustering of symptom trajectories based on DTW consisted of two steps. First, the DTW distance between each pair of symptom trajectories was calculated. This is illustrated in Fig. 1 with the example of two HRSD-17 item time-series (item 1 "depressed mood" and item 7 "work and interest") of a single patient. The temporal scoring (per 2 weeks) on the given items is seen in Fig. 1a, with the two items for which the DTW distance is calculated shown in red. This patient had 14 assessments over a period of 26 weeks. The trajectories of items 1 "depressed mood" and 7 "work and interest" over time are plotted in Fig. 1b. The deformations of the time axes between both items are added, which brings the two time-series as close as possible to each other, in which all elements must be matched. Next, the calculation of the shortest path between the two time-series is shown in Fig. 1c. The two time-series were aligned in time with compressions and expansions. The "symmetricP0" step pattern was used as the dynamic time warping algorithm to match the two sequences, resulting in the red "warping path." A Sakoe-Chiba Band of 2 was used in order for the severity scores to be matched to a maximum of plus or minus two time points (plus or minus a maximum of 4 weeks). Resulting from the DTW method, a distance measure (d) is produced: items with the best alignment, having a more similar slope and other dynamics (i.e., changes that covary over time), resulted in the smallest distance. The distance measures of each of the 17 time-series of individual HRSD-17 items are grouped in a distance matrix, comprising $(17^2 - 17)/2 = 136$ distances for each individual patient.

Second, this matrix of 136 distances was presented in a heatmap and used in a hierarchical cluster analysis and a symptom network per patient. For the hierarchical cluster analysis, each item is initially assigned to its own cluster, and then the algorithm proceeds

iteratively, at each stage joining the two most similar clusters, continuing until there is just a single cluster. We assumed 3 clusters for each patient, for illustrative purposes only, to enable easier recognition of the symptom with the more similar trajectories. We excluded all symptoms with a score of 0 throughout follow-up, as these tended to cluster together most strongly as these symptom pairs will have a distance of 0. At each stage, distances between clusters are recomputed by the Lance-Williams dissimilarity update formula according to the "Ward.D2" clustering methods. With "Ward.D2," the total withincluster variance is minimized, and the dissimilarities are squared before cluster updating.

Using the "qgraph" package, the structure of the network based on the distance matrix was visualized per patient, providing another way of graphical presentation of the clusters. We followed the recommendations on network analysis written by the developers of the R package [39]. A network with up to 17 nodes (representing the individual HRSD-17 depression symptoms) is obtained and, connecting them, the edges representing the distances between symptom trajectories. The thickness of the edges indicates the strength of the longitudinal elastic covariation (thicker edges represent a shorter distance between the two symptom trajectories).

Inter-individual approach

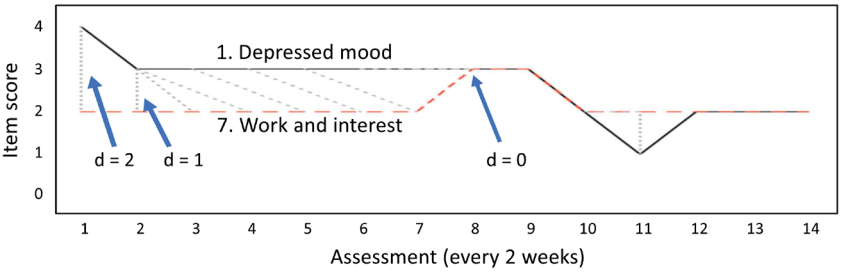
Next, we aimed to study the aggregated dynamics of individual symptoms to yield systematic patterns over time across patients. In this second part, we aimed to build a generalizable hierarchy of symptom clusters based on their shared temporal features. First, the 136 distances were averaged over the 255 patients, weighted for the number of assessments that were done for each of the patients (ranging from 2 through 17). Second, this matrix of 255 mean distances was used for the generalizable hierarchical cluster analysis. A scree plot was constructed displaying the heights in a downward curve and the elbow rule (i.e., the point where the graph leveled off) was used to determine the most appropriate number of clusters.

The "Distatis" algorithm from the "DistatisR" package was used to check whether using the actual 255 distance matrices instead of one mean distance matrix yielded similar clusters. Distatis is a generalization of classical multidimensional scaling (MDS), based on a three-way principal component analysis, to analyze a set of distance matrices. In order to compare these distance matrices, it combines them into a common structure called a compromise and then projects the original distance matrices onto this compromise. Compromise factors are calculated and plotted in the compromise space, with each component been given the length corresponding to its eigenvalues. We plotted each of the 17 HRSD symptoms on an X-Y plane according to their first and second compromise factor values.

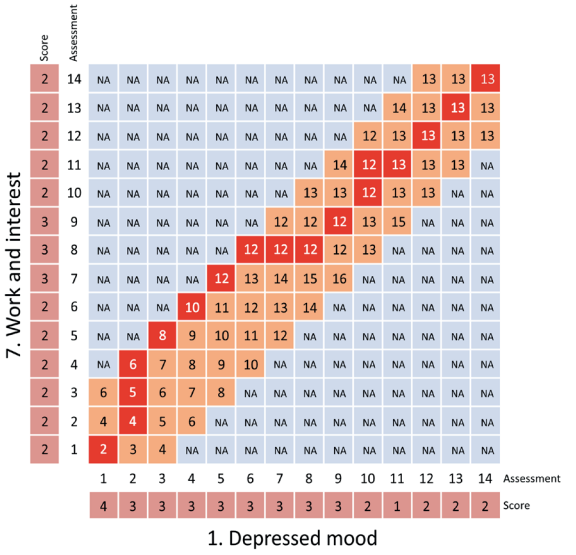
A. HRSD item score over time in one patient

Assessment (per 2 weeks)	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Depressed mood	4	3	3	3	3	3	3	3	3	2	1	2	2	2
2. Guilt	1	1	1	2	1	1	1	1	1	1	1	1	1	1
3. Suicide	1	1	0	1	1	1	1	1	1	0	0	1	1	1
4. Early insomnia	4	4	4	4	4	0	0	0	0	0	0	0	0	0
5. Middel insomnia	0	2	0	0	0	0	0	0	0	0	0	0	0	0
6. Late insomnia	4	2	4	4	4	2	2	2	2	0	0	0	2	0
7. Work and interests	2	2	2	2	2	2	2	3	3	2	2	2	2	2
8. Retardation	1	1	2	1	1	1	1	1	1	1	0	1	1	1
9. Agitation	0	1	0	1	1	1	1	1	0	0	0	0	0	1
10. Psychic anxiety	3	1	1	2	2	1	2	1	0	0	2	2	0	1
11. Somatic anxiety	1	0	0	1	1	0	0	1	0	0	0	1	0	0
12. Gastro-intestinal	2	2	2	2	0	2	2	0	0	0	0	0	0	0
13. General somatic	4	2	2	2	2	4	4	2	2	0	0	0	0	0
14. Genital symptoms	2	0	2	2	2	4	2	2	2	2	2	2	2	2
15. Hypochondriasis	0	0	1	0	0	0	0	0	0	0	0	0	0	0
16. Weight loss	1	1	0	0	0	0	0	0	0	0	0	0	0	0
17. Insight	0	0	0	0	0	0	0	0	0	0	0	0	0	0

B. Scores of items 1 and 7 over time, with DTW alignement



C. Optimal DTW warping route



D. Step pattern, contraint, and final distance

'symmetricP0' step pattern recursion:
g [i,j] = minimum of

- g [i-1 , j-1] + 2 * d [i , j]
- g [i , j-1] + d [i , j]
- g [i-1 , j] + d [i , j]

d = local distance
g = stretched distance

Sakoe Chiba Band:

- window size = 2
- constraints on the alignment, limiting the warping scope

DTW distance = 13
(according to the red warping curve)

Figure 1. For a single patient, the individual HRSD-17 item scores over time are shown (a). The DTW method uses a dynamic (i.e., stretching and compressing) programming approach to minimize a predefined distance measure (e.g., Euclidean distance), in order for the two time-series to become optimally aligned through a warping path (b). The optimal warping route between items 1 and 7 is shown (c). Using the "symmetricP0" step pattern and a Sakoe-Chiba Band of 2, this yields a final DTW distance of 13 (d)

In the following step, we investigated two centrality metrics, being closeness centrality and degree centrality [40] for the average distance matrix. Degree centrality is based on the number and strengths of connections each symptom has. Closeness centrality also takes the global network structure into account because it measures the average distance of a certain symptom to all other symptoms. Applied on the DTW data, closeness is inversely proportional to the mean DTW to all other symptoms and, in this way, indicates which symptom trajectory is the most similar to that of other symptoms. Finally, we computed the average DTW distance among all symptom trajectories for each patient. Symptoms that scored consistently zero were deleted from these analyses for that particular patient, as all such symptoms would result in distances of zero. Shorter average DTW distances reflected denser interconnections between symptoms, and longer average DTW distances reflected looser longitudinal connectivity between symptoms. In order to investigate the relationship between network density and reaching response and remission, we calculated the residuals of the regression with the number of assessments and the HDRS sum score. These residuals were plotted using box plots according to whether response and remission were reached, and we performed Wilcoxon signed-rank tests to compare the two samples. The analyses used the packages "dtw" (version 1.20.1), "pheatmap" (version 1.0.12), "parallelDist" (version 0.2.4), "qgraph" (version 1.6.2), and "DistatisR" (version) for the R statistical software (R version 3.6.0; R Foundation for Statistical Computing, Vienna, Austria, 2016). A sample code (with data from the of 2 exemplar patients of Fig. 2) can be found in Additional file 1.

Results

Patient characteristics

Table 1 shows the demographic and clinical characteristics and the use of psychotropic medication of the included patients. Patients had a mean age of 50.9 years (standard deviation [SD] = 15.4), and 165 were women (64.7%). A bipolar disorder was diagnosed in 48 patients (18.8%). The mean duration of illness was 11.2 ± 15 years. For 56 patients (22%), the current episode was the first depressive episode. The baseline HRSD-17 score was 20.7 (SD 4.6) on average, and 79.6% of the patients used antidepressants. Of the 255 patients, 169 showed treatment response and 128 remission at the end of admission. The median duration of hospitalization was 11 weeks, and the total number of assessments was 1480, with a mean of 5.8 and a median of 5 HRSD-17 assessments per patient.

Table 1. Characteristics and medication use of 255 consecutive depressed inpatients

Variable	Mean (SD) / No (%)
Demographic characteristics	
Female (%)	165 (64.7)
Bipolar disorder	48 (18.8)
Age in years, mean (SD)	50.9 ± 15.4
Education:[a,b] (n, %)	
- Lower	41 (16.1)
- Intermediate	119 (46.7)
- Higher	93 (36.5)
Work status: (n, %)	
- Unemployed	149 (58.4)
- Employed	93 (36.5)
- Other (student/voluntary service)	13 (5.1)
Marital status: (n, %)	
- Married	109 (42.7)
- Divorced/widowed	65 (25.5)
- Never married	81 (31.8)
Living situation:[c] (n, %)	
- Living alone	76 (29.8)
- Living with partner	96 (37.6)
- Living with family	83 (32.5)
Clinical characteristics	
Index (first) depressive episode (n, %)	56 (22)
History of 4 or more depressive episodes (n, %)	76 (29.8)
Lifetime substance abuse/dependency:[d]	
- Alcohol	27 (10.6)
- Drugs (THC, hard drugs, benzodiazepines)	5 (2.0)
Melancholic features	159 (62.4)
ROM baseline total scores:	
- HRSD-17	20.7 ± 4.6
- BDI-II	33.7 ± 9.2
Baseline medication use:	
- Antidepressants	203 (79.6)
- Antipsychotics	118 (46.3)
- Mood stabilizers	34 (13.3)
Responders (%)	169 (66.3)
Remitters (%)	128 (50.2)

Data are mean (SD) or No (%), when appropriate. ROM: routine outcome monitoring, HRSD-17: Hamilton Depression Rating Scale, BDI-II: Beck Depression Inventory.

[a] Lower education: general basic education only; intermediate education: middle vocational education; higher education: higher vocational education or university.

[b] Two missing values for education.

[c] Living alone includes living in a home for elderly and the convent

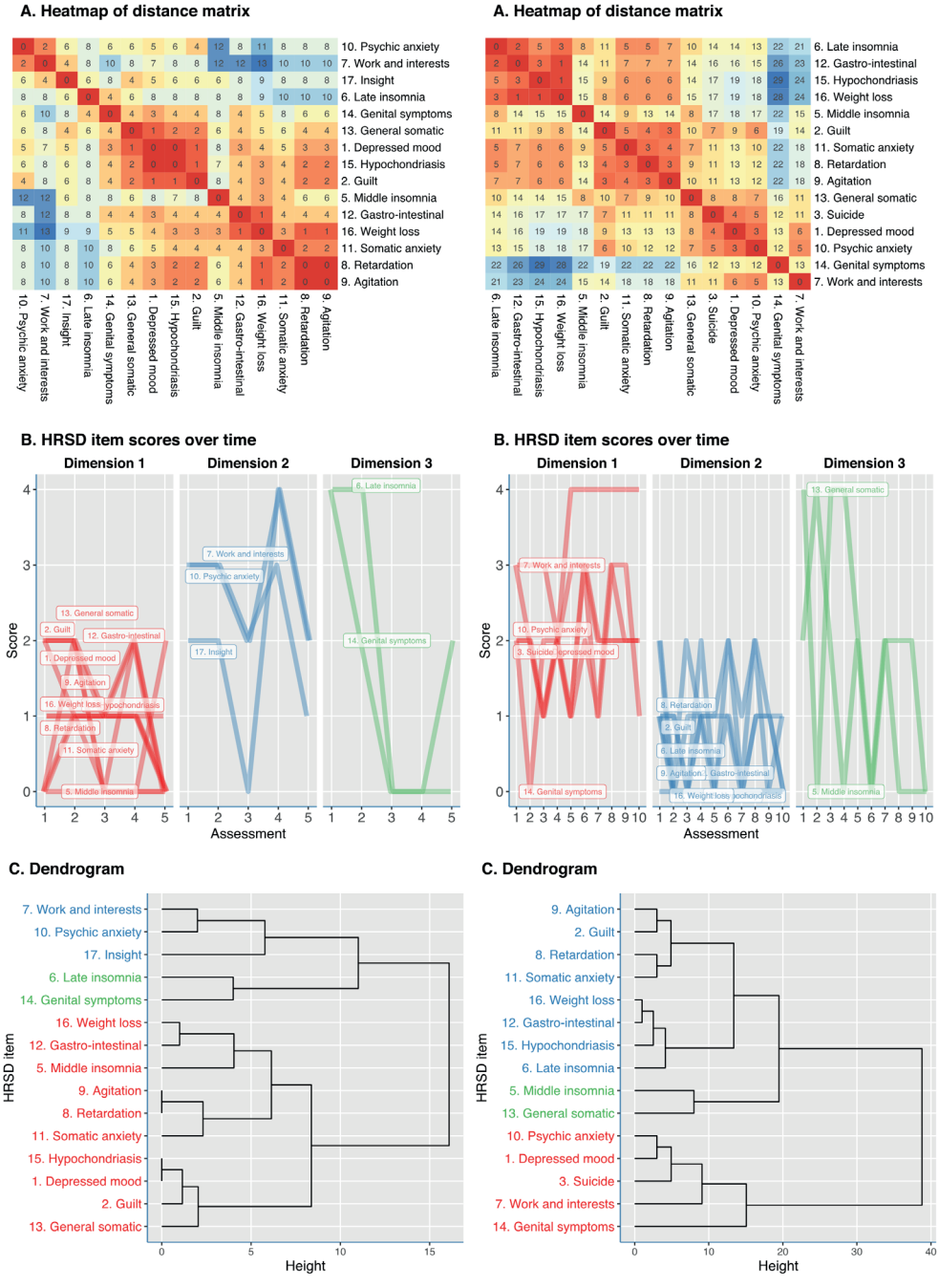
[d] Investigated using the MINI modules on substance abuse and suicidality

Intra-individual approach

In Fig. 2, the DTW analyses of 2 exemplar patients are shown. We will discuss these two exemplar patients in order to demonstrate the opportunities of the DTW clustering method to inform clinical practice. By comparing the results from patients 196 and 201, we can already see a high degree of inter-individual variability in symptom trajectories.

Patient 196 was a 55-year-old female presenting with psychotic depression. At admission, anhedonia, insomnia, and psychic anxiety were overtly present. The anxious preoccupations disabled her in engaging any psychotherapy program at the start of hospitalization. A treatment with electroconvulsive therapy (ECT) led to a resolution of the most central symptoms (symptom with most dense connections with other symptoms, e.g., depressed mood, feelings of guilt, and somatic anxiety) during the hospitalization of 2 months. Although she remained to score relatively high on the HRSD-17 symptoms "work and interests," "psychic anxiety," and "insight," she could be discharged after the resolution of the majority of her depressive symptoms. The central (red) symptoms tended to fluctuate most strongly together over time. Furthermore, due to the presence of residual symptoms that were resistant to ECT treatment, we formulated an advice for ambulatory psychological therapy to focus on these persistent (blue) symptoms of insight, engagement in activities, and psychic anxiety as a cornerstone of further treatment.

Patient 201 was a 38-year-old female who presented with a severely depressed mood and suicidal thoughts. She described her depressive complaints as an overpowering sense of feeling down and agitated. There was no loss of appetite or weight loss. A treatment with nortriptyline and trazodone (for her sleeping problems, mainly middle insomnia) was started. The sleeping problems improved quickly. Her depressed mood and suicidal thoughts did not change at the beginning of treatment. Treatment with lithium, because of a suspicion of underlying bipolar disorder, led to a quick resolution of the mood and anxiety symptoms. Loss of sexual interest was initially not present but commenced during hospitalization, possibly as a side effect of treatment.


A. Heatmap of distance matrix
10. Psychic anxiety
7. Work and interests
17. Insight
6. Late insomnia
14. Genital symptoms
13. General somatic
1. Depressed mood
15. Hypochondriasis
2. Guilt
5. Middle insomnia
12. Gastro-intestinal
16. Weight loss
11. Somatic anxiety
8. Retardation
9. Agitation
A. Heatmap of distance matrix
6. Late insomnia
12. Gastro-intestinal
15. Hypochondriasis
16. Weight loss
2. Guilt
11. Somatic anxiety
8. Retardation
9. Agitation
13. General somatic
3. Suicide
1. Depressed mood
10. Psychic anxiety
14. Genital symptoms
7. Work and interests
B. HRSD item scores over time
Dimension 1
Dimension 2
Dimension 3
Score
Assessment
B. HRSD item scores over time
Dimension 1
Dimension 2
Dimension 3
Score
Assessment
C. Dendrogram
HRSD item
Height
C. Dendrogram
HRSD item
Height

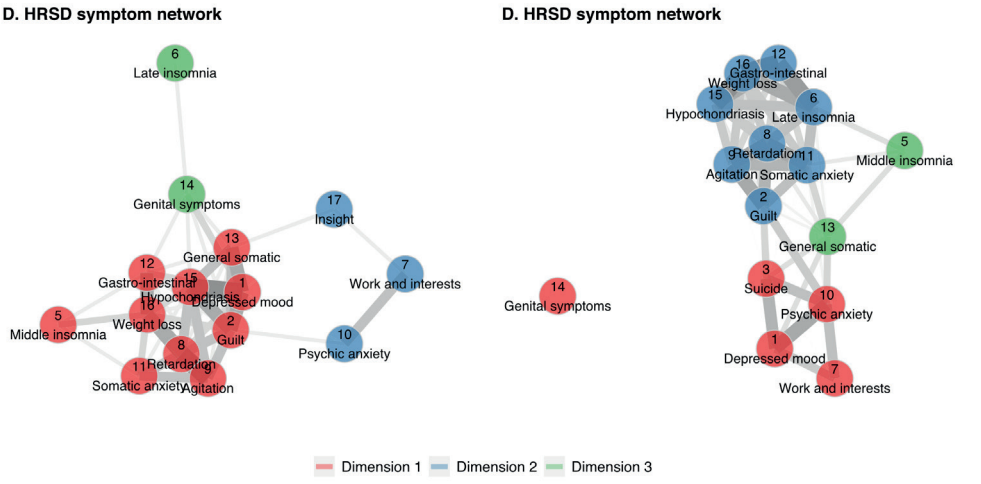


Figure 2. DTW analysis of HRSD-17 symptoms for patient no. 196 and patient no. 201 (a). Heatmap (symptoms that show high correlation are given a "hot" red color, and those that are not correlated are given a "cold" blue color) (b). Dendrogram based on the clustering of DTW distances of 15 of the non-zero HRSD-17 item scores over time (c). Network graph based on the distance matrix: connections between symptoms (edges) indicate distances between symptom trajectories (d). Centrality statistics of the network graph: centrality is based on the number and strengths of connections each symptom has. Closeness also takes the global network structure into account (e)

Inter-individual approach

Figure 3 shows the nomothetic analysis of the 255 patients. A total of five clusters emerged, based on the elbow method in the scree plot (see Fig. 3a). The hierarchical cluster analysis was estimated based on the average weighted distance matrix (Fig. 3b). These clusters consisted of symptoms with a similar course trajectory: (1) core symptoms (2 items: "depressed mood,""work and interests"), (2) sleep symptoms (3 items: late, middle, and early insomnia), (3) distress (2 items: "guilt," "psychic anxiety"), (4) somatic symptoms (2 items: "genital symptoms," "general somatic symptoms"), and (5) inner turmoil (8 items: insight, weight loss, hypochondriasis, gastro-intestinal symptoms, somatic anxiety, agitation, retardation, and suicide). As is shown in Additional file 2: Fig. S1, the network plots did not change significantly when excluding all symptoms with a score of 0 throughout follow-up.

In the following step, we analyzed the actual 255 distance matrices, instead of one mean distance matrix, using Distatis. Figure 3c shows the Distatis compromise plot in which each HRSD-17 item is plotted according to their first and second compromise factor values. The distribution pattern of the HRSD-17 items in the compromise plot shows a comparable pattern to the obtained hierarchical clusters, corroborating the obtained five clusters. The clusters corroborated those found with the hierarchical cluster analysis on the average distance matrix. Next, the average distance matrix was visually presented

in a network graph in Fig. 3d. Figure 4 shows the two centrality measures based on the network from the average distance matrix. These can inform us on which symptoms globally tend to covary together over time. The items from the "inner turmoil" show the highest degree centrality and closeness centrality scores, indicating that they covaried most strongly with other HRSD-17 items. Items constituting the "insomnia" or "somatic symptom" cluster showed lower centrality which suggests that these symptoms behaved in a more independent manner over time compared to the other HRSD-17 items.

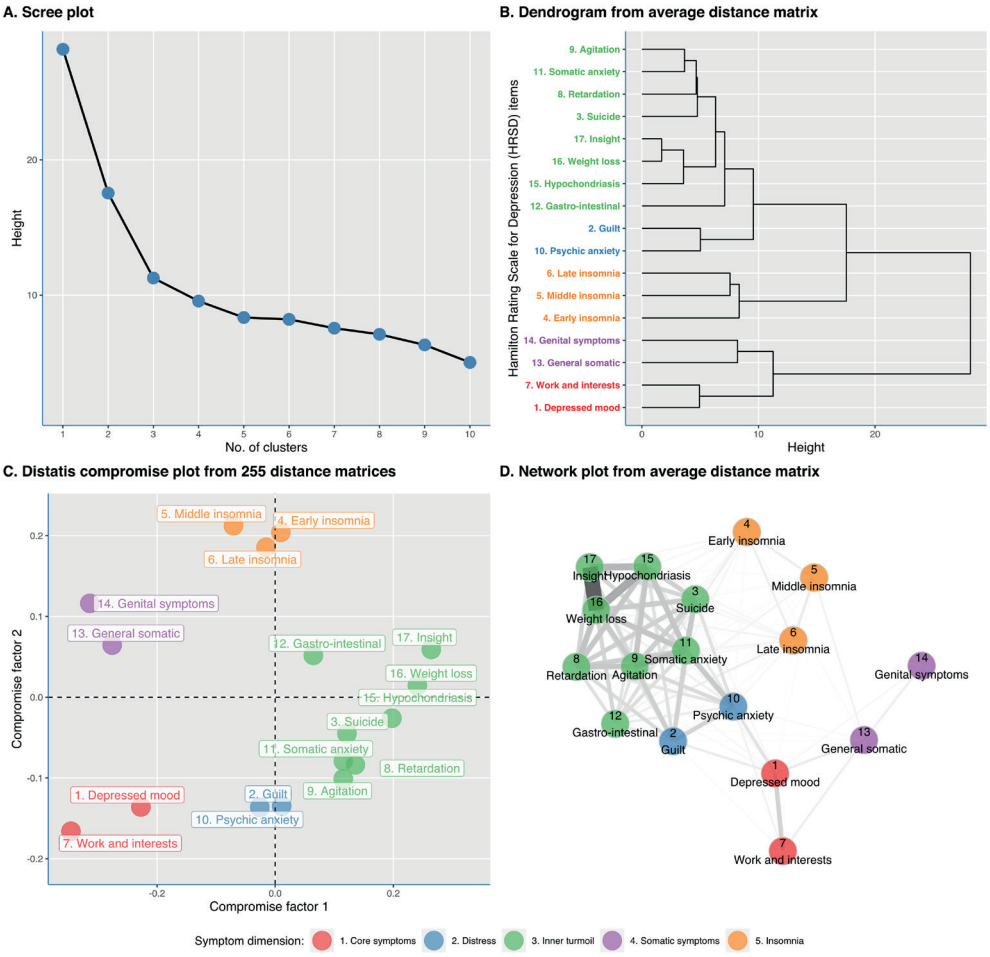


Figure 3. Nomothetic analyses based on all distance matrices from 255 depressed inpatients. The scree plot displays the eigenvalues in a downward curve. The number of factors was determined using the elbow method (i.e., the point where the slope of the curve is leveling of; in our example, this is 5: after this point, the slope of the curve is nearly stable) (a). Ward's (D2, i.e., general agglomerative hierarchical clustering procedure) clustering criterion on the weighted mean distance matrix from 255 patients (b). Distatis analysis: the PCA of the compromise matrix (i.e., weighted average of individual cross-product matrices) gives the position of the objects in the compromise space (c). Overview of the networks of HRSD-17 items for 255 patients (d)

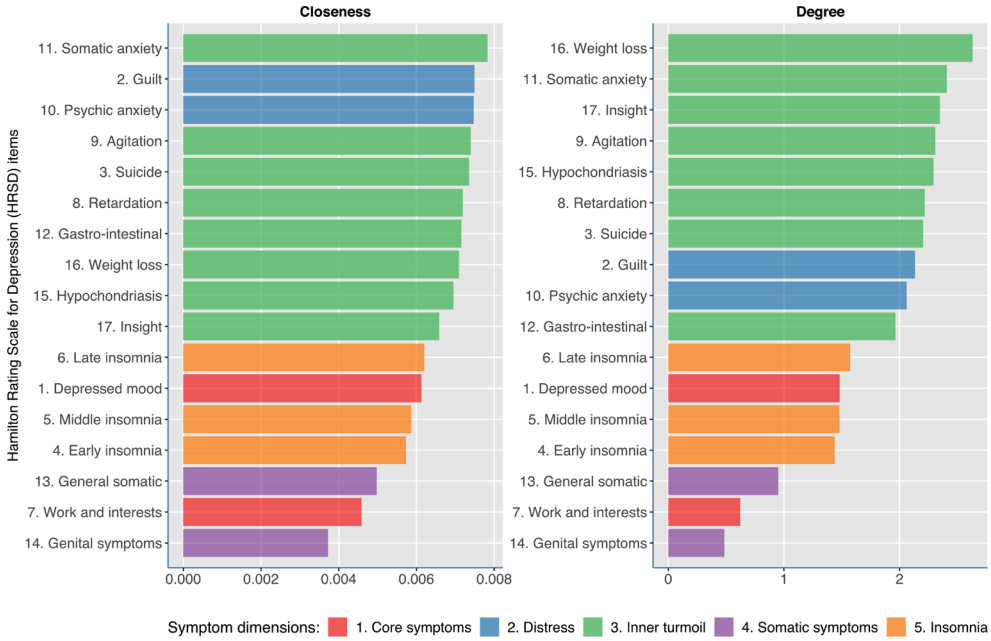


Figure 4. Centrality measures. The closeness centrality is the inverse of the average length of the shortest path between the focal node and every other node in the network (i.e., the more central a node is, the closer it is to all other nodes). Degree centrality represents the connectivity, based on the number and strengths of edges connected to it

The evolution of the mean HRSD-17 item levels over time are visualized in Fig. 5a using mixed models per item. The eight items with a range from 0 to 2 (three insomnia items: gastro-intestinal complaints, general somatic and genital symptoms, insight and weight loss) were scaled to a range from 0 to 4 in order to make a comparability between all items possible. The HRSD-17 items "depressed mood," "work and interests," "general somatic," and "genital symptoms" had the highest baseline severity. The items "insight" and "weight loss" had the lowest mean scores and stayed relatively low during hospitalization. Figure 5b shows the intercepts and slopes of the 17 mixed models for the individual longitudinal trajectories. The intercepts indicated that genital, general somatic, and depressed mood symptoms generally scored the highest at baseline. The slopes of the linear model revealed that depressed mood showed the steepest decline over time.

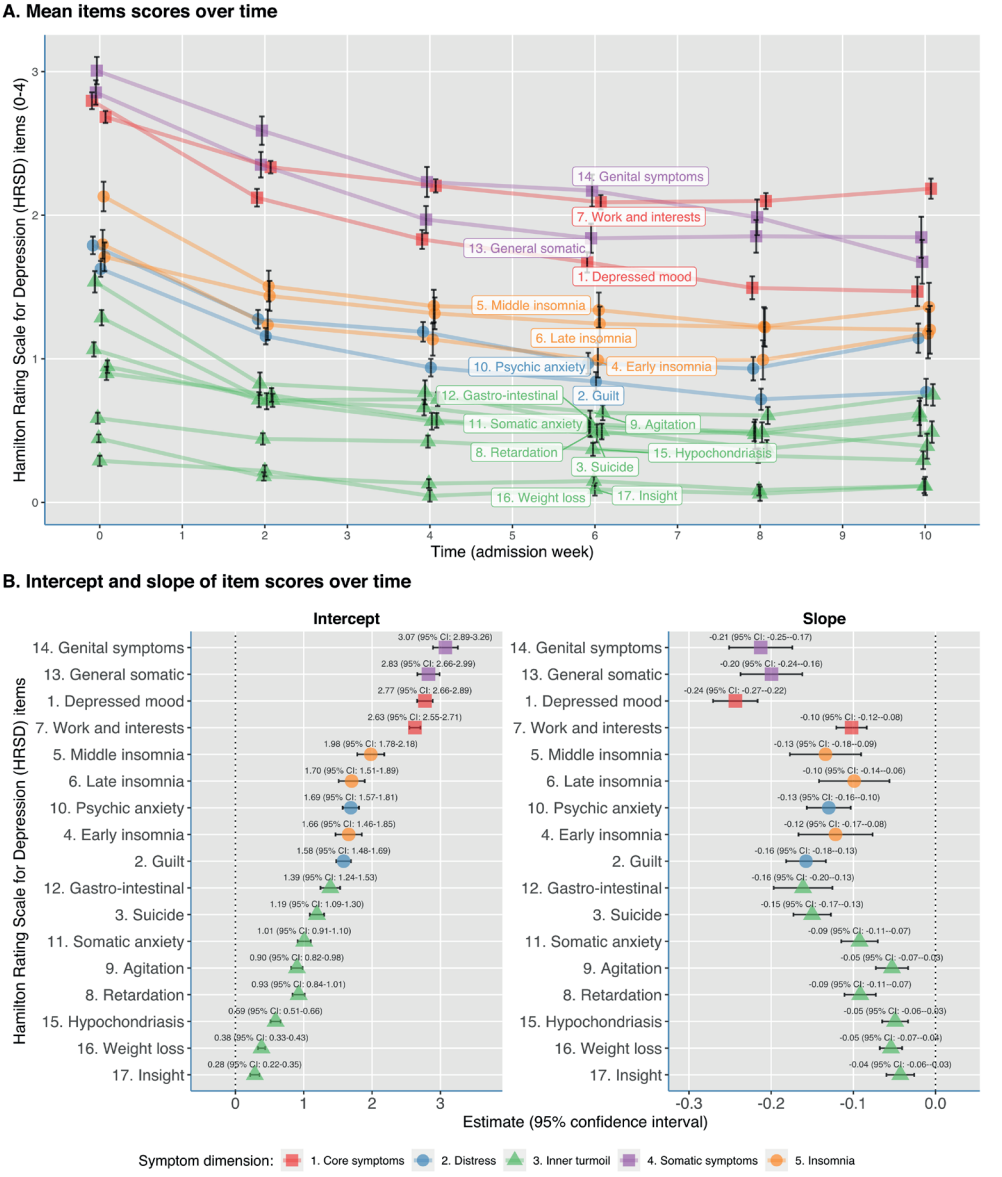


Figure. 5. Forest plot of the 17 HRSD items of two mean levels of indicators of individual longitudinal trajectories. The mixed model intercept (a) indicates which symptoms scored the highest at baseline (i.e., genital, general somatic symptoms, and depressed mood scored the highest atbaseline). The mixed model slope (b) of the linear model showed the average decline per 2-week time interval (i.e., depressed mood showed the steepest decline over time)

As shown in Fig. 6, patients that reached response or remission during hospitalization had significantly shorter average distances among symptoms than patients who failed to reach response or remission. That is, patients reaching response or remission mostly had on average a more densely connected symptom network (based on the mean DTW analysis). We excluded symptoms that scored zero throughout the admission, yet when these symptoms were included, this resulted in similar findings (see Additional file 2: Fig. S2A). In addition, not adjusting for HRSD-17 total scores at baseline did not alter the results (see Additional file 2: Fig. S2B). Exploring the difference in network connectivity between unipolar and bipolar depressed patients revealed a denser symptom network in bipolar than in unipolar patients (see Additional file 2: Fig. S3).

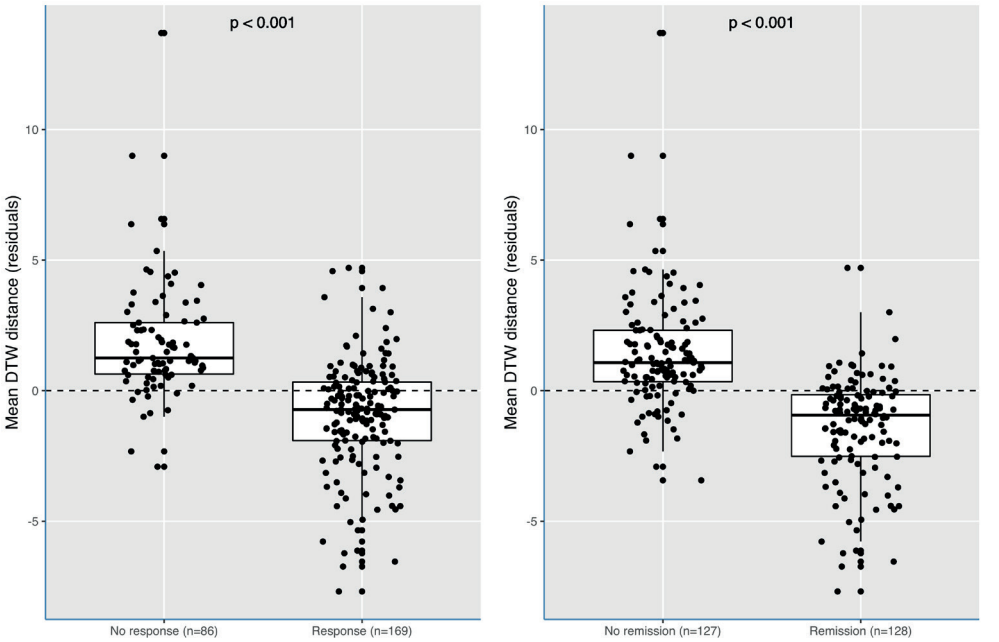


Figure. 6 Average DTW distance according to response and remission. Those patients with response or remission had the shortest average distance among symptom trajectories, indicating denser interconnections (p by Wilcoxon signed-rank test to compare the two samples). Distances were adjusted for the number of assessments and the baseline total HRSD-17 score

Discussion

The present study is the first to analyze the time-series of depression symptoms using DTW analyses in psychiatric inpatients. We applied the DTW computational method to estimate and visualize similarities in symptom trajectories and to yield clusters of symptoms with similar course trajectories both at the patient level and at the group level. Both the intra- and inter-individual analyses may help to increase our insight into the dynamical complexity of symptom trajectories in severely depressed inpatients. Furthermore, combining ROM techniques with automated feedback for the clinician based on the methods—as introduced here—proved useful to inform and facilitate clinical decision making. Overall, three major findings are worth discussing in more detail.

We first focused on the individual symptom dynamics that proved to be highly variable across individuals and thus idiosyncratic [41]. This finding supports previous concerns on the use of sum scores for assessing treatment outcome, since sum scores do not represent this dynamical symptom complexity well [4, 16], with a loss of substantial information that may be of clinical relevance. The intra-individual dynamic symptom clusters and symptom networks, in which the edges between symptoms represented the dynamical relation between them, allowed us to gain insight into the relative importance of certain symptoms for individual patients, but also at the group level [40]. Such symptoms may cause other symptoms, which may be different for other patients [9]. It could be hypothesized that targeting treatment on such central symptoms early in therapy may lead to a more rapid resolution of closely connected depressive symptoms [42, 43].

Overall, the study of intra individual temporal dynamics of depression symptoms is rare in the literature. A growing field of research has focused on the development of individual dynamic networks of symptoms in which time-series or experience sampling methods (ESM) data are used to study the within-patient dynamical structure of symptoms [43–47]. These networks are mostly estimated using vector autoregression (VAR) which estimates both lagged (i.e., time minus one temporal) and contemporaneous (i.e., simultaneous) relationships among multiple symptoms [45]. The DTW approach represents a less constraint analysis of individual symptom dynamics since the DTW distance measure accounts for a longer time period when measuring the similarity between each pair of depressive symptoms (2 time points). Furthermore, it provides an accessible and easily interpretable method that can be a useful tool for clinicians and researchers for the early detection of central symptoms and the directed tailoring of treatment towards these symptoms. Moreover, it does not necessitate the time-consuming ESM data collection, which can be challenging in daily clinical practice due to the considerable burden it puts on participants.

Secondly, we focused on the group-level analyses, which yielded five symptom clusters with more similar dynamics over time across the total sample. Compared to cross-sectional factor analytic studies, which investigate the co-occurrence of depressive symptoms at a certain time point, we similarly found that the three sleep items appeared consistently in one factor [20, 21]. Previously, there was some support for the presence of a "general depression," with depressed mood, guilt, suicide, work and interests, and psychic anxiety appearing on one factor [20–22]. We, however, found that only "depressed mood" and "work and interest" showed the most consistent trajectories over time, which represented the core symptoms of depression. Somatic symptoms did not appear on the same factor as described for cross-sectional factor analyses ("somatic symptom" or "somatized depression" consisting of somatic symptoms, weight loss/gastro intestinal symptoms, loss of libido/genital symptoms) [20, 22]. Moreover, previous studies found evidence for limited longitudinal invariance, where the number of factors did not hold across time [24, 48] which is supported by our and previous data that the Omegas and Cronbach's alphas were not stable over time, but improved during hospital admission [49]. An internal validation of our findings using a random sample of 128 and 127 patients of our sample revealed the same dynamical clusters (see Additional file 2: Fig. S4). Further validation of these findings in an independent sample is necessary.

Third, we found that patients who reached response or remission during hospitalization had a more densely connected symptom network compared to patients that failed to reach response or remission. This contrasts with the findings of Van Borkulo et al. [50], who identified a more densely connected cross-sectional depression symptom network in not remitters compared to patients reaching remission. Although the method of quantifying network connectivity was not the same (average DTW distance versus network comparison test), this shows, once again, the importance of studying longitudinal networks besides cross-sectional symptom networks. Our findings could be related to the literature reviewed by Scheffer [51], showing that networks with high connectivity can change more abruptly (for better and worse) in response to external events (so-called critical transitions). Applied on the DTW network analysis, when symptoms have a low level of connectivity, they seem to behave more independently from each other and in response to an external factor such as admission and treatment, which may have lowered the probability of an acute response or remission to treatment. These findings need to be confirmed in further studies, as the definition of response and remission was also based on the HRSD sum scores, which is not independent of the DTW assessments from HRSD time-series data.

The DTW method has a promising potential for clinical practice, and it builds further upon the already available evidence of the value of measurement-based care in psychiatry [52]. First, the DTW symptom clusters allow the clinician to gain insight into the dynamics of individual depression symptoms and longitudinal symptom clusters. Second, as illustrated

in the two idiographic analyses (Fig. 2), the DTW method has the potential to facilitate clinical decision making. More specifically, treatment interventions targeted at the most central symptom (i.e., symptoms with the most dense connections with other symptoms) could lead to a rapid resolution of the depressive syndrome. Third, the graphical representation of the DTW clusters is easily amenable as a feedback tool for patients to gain more insight into the central symptoms that tend to covariate with a variety of other symptoms or in symptom clusters that tend to move in a more independent matter. This could lead to a more nuanced insight in reaching response or remission or lack thereof.

An important strength of our study is the use of the innovative DTW clustering method to study the timeseries of individual symptom severity scores. The DTW method is able to process the highly dimensional ROM time-series data in order to reduce the complexity of the data while still maintaining the essential characteristics of the dataset. By using an elastic measurement, DTW provides an optimal time alignment between two timeseries. Furthermore, DTW can be accurately used in smaller datasets and individual patients [31]. Another strength of our study is the relatively complete dataset of ROM data from real-world consecutive inpatients. Nonetheless, our results must be considered in light of some limitations. First, exclusively inpatients were recruited from one center which may limit the generalizability of our results to outpatients and other patient groups. Second, patients were treated with a variety of different combinations of psychotropic drugs which likely affected the course and dynamical characteristics of individual symptoms (such as concentration difficulties in those receiving ECT). Future studies using data from randomized trials may help to unravel the influence of different treatment strategies on the dynamic symptom dimensions. Third, the HRSD-17 is not designed to investigate individual symptoms, and its items are scored on a crude scale with only three or five answer categories resulting in low variability and precision. Fourth, assessments were done with 2-week intervals, and DTW analyses may be more useful in more frequent time-series like those collected with ESM. Fifth, the DTW method allows some flexibility in how it is applied to study MDD symptom trajectories, e.g., in terms of the global constraint (Sakoe-Chiba Band). We adopted default settings based on simulation studies in the prior literature and hope that future methodological studies working with psychiatric data specifically will investigate how robust empirical results are to changes in default settings of the DTW method, e.g., using multiverse analyses [53].

Conclusion

MDD is a heterogeneous disorder consisting of dynamic symptom clusters that varied between patients. The use of repeated, standardized clinical rating scales yields extensive information on patient-specific symptoms dynamics. DTW may be a promising new

methodology for the study of the complex dynamic system of interacting psychiatric symptoms [9, 15, 54] with the potential to facilitate personalized psychiatry care.

Supplementary Material

The online version contains supplementary material available at https://doi.org/10.1186/s12916-020-01867-5.

References

1. Borsboom D. Psychometric perspectives on diagnostic systems. J Clin Psychol. 2008 Sep;64(9):1089–108.
2. Schmittmann VD, Cramer AOJ, Waldorp LJ, Epskamp S, Kievit RA, Borsboom D. Deconstructing the construct: a network perspective on psychological phenomena. New Ideas Psychol. 2013;31(1):43–53.
3. Borsboom D, Mellenbergh GJ, Van Heerden J. The theoretical status of latent variables. Psychol Rev. 2003;110:203–19.
4. Fried EI. Problematic assumptions have slowed down depression research: why symptoms, not syndromes are the way forward. Front Psychol. 2015;6:309.
5. Fried EI, Nesse RM. Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. BMC Med. 2015;13(1):72
6. Trajković G, Starčević V, Latas M, Leštarević M, Ille T, Bukumirić Z, et al. Reliability of the Hamilton Rating Scale for Depression: a meta-analysis over a period of 49 years. Psychiatry Res. 2011;189(1):1–9.
7. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry. 1979;134(4):382–9
8. Gibbons RD, Clark DC, Kupfer DJ. Exactly what does the Hamilton depression rating scale measure? J Psychiatr Res. 1993;27(3):259–73.
9. Borsboom D, Cramer AOJ. Network analysis: an integrative approach to the structure of psychopathology. Annu Rev Clin Psychol. 2013;9(1):91–121.
10. Cramer AOJ, Waldorp LJ, Van Der Maas HLJ, Borsboom D. Comorbidity: a network perspective. Behav Brain Sci. 2010;33:137–50.
11. Hasler G, Northoff G. Discovering imaging endophenotypes for major depression. Mol Psychiatry. 2011;16(6):604–19
12. Myung W, Song J, Lim SW, Won HH, Kim S, Lee Y, et al. Genetic association study of individual symptoms in depression. Psychiatry Res. 2012;198(3):400–6.
13. Kendler KS, Aggen SH, Neale MC. Evidence for multiple genetic factors underlying DSM-IV criteria for major depression. JAMA Psychiatry. 2013;70(6): 599–607.
14. Faravelli C, Servi P, Arends JA, Strik WK. Number of symptoms, quantification, and qualification of depression. Compr Psychiatry. 37(5):307–15
15. Kendler KS, Zachar P, Craver C. What kinds of things are psychiatric disorders? Psychol Med. 2011;41(6):1143–50
16. Fried EI, Nesse RM, Zivin K, Guille C, Sen S. Depression is more than the sum score of its parts: individual DSM symptoms have different risk factors. Psychol Med. 2014;44(10):2067–76
17. Beltz AM, Wright AGC, Sprague BN, PCM M. Bridging the nomothetic and idiographic approaches to the analysis of clinical data. Assessment. 2016; 23(4):447–58

18. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research Domain Criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry. 2010;167:748–51.
19. Insel TR. The NIMH Research Domain Criteria (RDoC) project: precision medicine for psychiatry. Am J Psychiatry. 2014;171:395–7.
20. Shafer AB. Meta-analysis of the factor structures of four depression questionnaires: Beck, CES-D, Hamilton, and Zung. J Clin Psychol. 2006;62(1): 123–46
21. Bagby RM, Ryder AG, Schuller DR, Marshall MB. The Hamilton Depression Rating Scale: has the gold standard become a lead weight? Am J Psychiatry. 2004;161:2163–77.
22. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23(1):56–62
23. Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol. 1967;6(4):278–96.
24. Fried EI. Are more responsive depression scales really superior depression scales? J Clin Epidemiol. 2016;77:4–6.
25. Kasper S, Dienel A. Cluster analysis of symptoms during antidepressant treatment with Hypericum extract in mildly to moderately depressed outpatients. A meta-analysis of data from three randomized, placebo-controlled trials. Psychopharmacology (Berl). 2002;164(3):301–8
26. Chekroud AM, Gueorguieva R, Krumholz HM, Trivedi MH, Krystal JH, Mccarthy G. Reevaluating the efficacy and predictability of antidepressant treatments a symptom clustering approach supplemental content. JAMA Psychiatry. 2017;74(4):370–8
27. de Beurs E, den Hollander-Gijsman ME, van Rood YR, van der Wee NJ, Giltay EJ, van Noorden MS, van der Lem R, van Fenema E, Zitman FG. Routine outcome monitoring in the Netherlands: practical experiences with a webbased strategy for the assessment of treatment outcome in clinical practice. Clin Psychol Psychother. 2011;18(1):1-12.
28. Washington AE, Lipstein SH. The Patient-Centered Outcomes Research Institute—promoting better information, decisions, and health. N Engl J Med. 2011;365(15):e31
29. Sakoe H, Chiba S. Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans Acoust. 1978;26(1):43–9.
30. Introduction. In: Information retrieval for music and motion. Berlin, Heidelberg: Springer Berlin Heidelberg; 2007. p. 1–13. Available from: http://link.springer.com/10.1007/978-3-540-74048-3_1.
31. Ding H, Trajcevski G, Scheuermann P, Wang X, Keogh E. Querying and mining of time series data. Proc VLDB Endow. 2008;1(2):1542–52.
32. Dandu SR, Engelhard MM, Qureshi A, Gong J, Lach JC, Brandt-Pearce M, et al. Understanding the physiological significance of four inertial gait features in multiple sclerosis. IEEE J Biomed Heal Informatics. 2018;22(1):40–6
33. Engelhard M, Dandu SR, Lach J, Goldman M, Patek S. Toward detection and monitoring of gait pathology using inertial sensors under rotation, scale, and offset invariant dynamic time warping. In: Proceedings of the 10th EAI International Conference on Body Area Networks.

ICST; 2015. Available from: http://eudl.eu/doi/10.4108/eai.28-9-2015.2261503. [cited 2020 Mar 30].

34. Li M, Tian S, Sun L, Chen X. Gait analysis for post-stroke hemiparetic patient by multi-features fusion method. Sensors (Basel). 2019;19(7):1737.
35. Zhang G, Kinsner W, Huang B. Electrocardiogram data mining based on frame classification by dynamic time warping matching. Comput Methods Biomech Biomed Engin. 2009;12(6):701–7 Available from: http://www.ncbi.nlm.nih.gov/pubmed/19360509.
36. Csikszentmihalyi M, Larson R. Validity and reliability of the experiencesampling method. J Nerv Ment Dis. 1987;175(9):526–36.
37. Fisher AJ, Newman MG, Molenaar PCM. A quantitative method for the analysis of nomothetic relationships between idiographic structures: dynamic patterns create attractor states for sustained posttreatment change. J Consult Clin Psychol. 2011;79(4):552–63.
38. Hebbrecht K, Stuivenga M, Birkenhäger T, Van Der Mast RC, Sabbe B, Giltay EJ. Symptom profile and clinical course of inpatients with unipolar versus bipolar depression. Neuropsychobiology. 2019;79:4–5
39. Epskamp S, Borsboom D, Fried EI. Estimating psychological networks and their accuracy: a tutorial paper. Behav Res Methods. 2018;50(1):195–212.
40. Opsahl T, Agneessens F, Skvoretz J. Node centrality in weighted networks: generalizing degree and shortest paths. Soc Networks. 2010; 32(3):245–51.
41. Fisher AJ. Toward a dynamic model of psychological assessment: implications for personalized care. J Consult Clin Psychol. 2015;83(4):825–36.
42. Fried EI, Epskamp S, Nesse RM, Tuerlinckx F, Borsboom D. What are "good"depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. J Affect Disord. 2016;189:314–20.
43. Epskamp S, van Borkulo CD, van der Veen DC, Servaas MN, Isvoranu AM, Riese H, et al. Personalized network modeling in psychopathology: the importance of contemporaneous and temporal connections. Clin Psychol Sci. 2018;6(3):416–27
44. Beltz AM, Wright AGC, Sprague BN, PCM M. Bridging the nomothetic and idiographic approaches to the analysis of clinical data. Assessment. 2016; 23(4):447–58
45. Bringmann LF, Ferrer E, Hamaker EL, Borsboom D, Tuerlinckx F. Modeling nonstationary emotion dynamics in dyads using a time-varying vectorautoregressive model. Multivariate Behav Res. 2018;53(3):293–314.
46. Bulteel K, Tuerlinckx F, Brose A, Ceulemans E. Improved insight into and prediction of network dynamics by combining VAR and dimension reduction. Multivariate Behav Res. 2018;53(6):853–75.
47. Fisher AJ, Reeves JW, Lawyer G, Medaglia JD, Rubel JA. Exploring the idiographic dynamics of mood and anxiety via network analysis. J Abnorm Psychol. 2017;126(8):1044–56.
48. Steinmeyer EM, Möller HJ. Facet theoretic analysis of the Hamilton-D scale. J Affect Disord. 1992;25(1):53–61.

49. Fried EI, van Borkulo CD, Epskamp S, Schoevers RA, Tuerlinckx F, Borsboom D. Measuring depression over time...or not? lack of unidimensionality and longitudinal measurement invariance in four common rating scales of depression. Psychol Assess. 2016;28(11):1354–67.
50. Van Borkulo C, Boschloo L, Borsboom D, Penninx BWJH, Lourens JW, Schoevers RA. Association of symptom network structure with the course of longitudinal depression. JAMA Psychiatry. 2015;72(12):1219–26.
51. Van De Leemput IA, Wichers M, Cramer AOJ, Borsboom D, Tuerlinckx F, Kuppens P, et al. Critical slowing down as early warning for the onset and termination of depression. Proc Natl Acad Sci U S A. 2014;111(1):87–92.
52. Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, et al. Evaluation of outcomes with citalopram for depression using measurementbased care in STAR*D: implications for clinical practice. Am J Psychiatry. 2006;163(1):28–40.
53. Steegen S, Tuerlinckx F, Gelman A, Vanpaemel W. Increasing transparency through a multiverse analysis. Perspect Psychol Sci. 2016;11(5):702–12.
54. Cramer AO, van Borkulo CD, Giltay EJ, van der Maas HL, Kendler KS, Scheffer M, Borsboom D. Major depression as a complex dynamic system. PLoS One. 2016;11(12):e0167490

CHAPTER 6

Insight into the directed symptom dynamics of depressed inpatients: a dynamic time warp analysis of the Beck Depression Inventory-II

Hebbrecht K, Fried EI, Stuivenga M, Birkenhäger T, Morrens M, Sabbe B, Giltay EJ.

Manuscript submitted

Abstract

Background: Depressive symptoms interact in complex ways within individual patients over time. We apply Dynamic Time Warp (DTW) to data of depressed inpatients, which provides insight into the dynamic relations among symptoms by assessing the similarity of changes over time between symptom pairs.

Methods: The 21-item BDI-II was assessed bi-weekly in 166 depressed inpatients, with on average 6.9 assessments per patient. DTW was used to analyze standardized symptom scores for each individual patient, yielding 166 symptom distance matrices. Undirected DTW analyses yielded groups of symptoms with a similar change profile over time (i.e., symptom dimensions) while directed DTW analyses identified which symptom changes temporally precede or follow other symptoms changes. Both DTW analyses were first done at the level of the individual patient and the data were subsequently aggregated to the group level.

Results: Patients were on average 50 years old and 63% were female. Undirected DTW analyses yielded four symptom dimensions: 'Core Depressive Symptoms' (6 items), 'Social Withdrawal and Lethargy' (7 items), 'Disturbed self-definition' (6 items), and Auto-aggression (2 items). Directed DTW analyses showed that worthlessness, pessimism, loss of interest 29 and tiredness had the highest outstrength centrality, despite the high between-person variability in individual-level analyses.

Conclusion: DTW promises to increase insight into the temporal dynamics of symptom changes in mood disorders. Symptoms with high outstrength, derived from directed DTW analyses, may have a positive feedback effect on other symptoms. Future studies are needed to investigate whether DTW analyses provide insights for improved prediction and treatment targeting in mood disorders.

Introduction

The latent disease model for psychopathology has dominated depression research for years. Under this framework, depression symptoms are considered as equal and independent consequences of a common cause: the underlying mental disorder [1]. Consequently, the vast majority of clinical research on major depressive disorder adds symptoms to one total score in order to measure change in severity over time [2]. However, this is at odds with findings from previous studies showing that individual depression symptoms have different risk factors [2], psychosocial impact [3], and underlying biology [4,5].

Network approaches offer a promising conceptual framework that is particularly suitable to study psychopathology as a complex dynamic system [6]. In such a system, symptom-symptom interactions evolve over time and give rise to emergent properties such as a depressive state or other mental disorders [7,8]. Symptom networks can be estimated using statistical network models, and visualized as a set of symptoms (or nodes) that are connected by edges that represent these statistical relationships between them (e.g., correlations or regression coefficients) [1,6]. Most network analyses have investigated cross-sectional relations between symptoms at a specific time point [9] mainly using partial correlation networks, directed acyclic graphs, and relative importance networks [10]. However, such studies do not capture the relation between symptoms over time [11,12]. Furthermore, findings from cross-sectional analyses may not generalize well from the group to the individual, where different people may exhibit profoundly different patterns of symptom relations. Therefore, there is growing recognition to first analyze symptom relations in individual patients separately (i.e., idiographic analyses) and subsequently aggregate these data to the group (i.e., nomothetic) level, instead of the other way around [13,14]. Using time-series data, the most used statistical network model to study dynamic symptom relations is based on lagged multilevel vector autoregressive (VAR) models [15]. Although this method has yielded many new insights, disadvantages of VAR models are the focus on a single time-interval as the unit of analysis and the assumption of stationarity. Further, VAR models require intensive longitudinal data with many dozens of measurement points, and severely depressed patients can be unable to complete so many measures. New analytical methods are needed when only a limited number of repeated measurements are available (i.e., panel data) [6,16].

In this study, we use Dynamic Time Warping (DTW) as an alternative analytic method for the analysis of longitudinal data, which is based on a shape-based distance measures. In two previously published studies, we have used DTW to obtain undirected symptom networks from the Hamilton Rating Scale for Depression (HRSD-17) [17] and the Comprehensive Psychopathological Rating Scale (CPRS) [18] in depressed patients. Undirected DTW analyses reveal which symptoms show a similar change profile over time (which we will

denote in this manuscript as symptom dynamics). However, such undirected analyses do not allow to disentangle whether changes in a symptom temporally precede or follow changes in another symptom. Since this is a necessary (albeit not sufficient step) towards causal inference and ultimately interventions, in the current study we aim to investigate both undirected as well as directed depression symptom dynamics (which investigate the directionality of symptom change over time).

We used 2-week intervals data over an admission period from 8 to maximum 34 weeks from the Beck Depression Inventory (BDI-II) [19], a widely used self-report scale that focusses on cognitive and affective symptoms, such as worthlessness and pessimism. It differs markedly from the previously analyzed HRSD-17 [17], which is a self-report scale developed for inpatients that relies heavily on clinical, observable signs [20]. The undirected DTW analysis yields symptom dimensions consisting of symptoms with similar dynamics over time. The directed DTW analysis, on the other hand, can identify whether changes in one symptom are temporally preceded or followed by other changes. The directed results may ultimately yield actionable insights for improved prediction and targets for treatment both at the individual as well as on the group level [16], in the spirit of precision and personalized medicine.

Methods

Sample and setting

Data were collected as part of the Hercules Routine Outcome Monitoring study [17]. The original study sample comprised 276 consecutive patients admitted to a tertiary psychiatric hospital (Duffel, Belgium) with an MDD or depressive episode in bipolar disorder (confirmed by the MINI-Plus Interview, based on the DSM-IV criteria). Exclusion criteria were restricted to maximally reflect the 'real-world' clinical depressed population; only patients with (comorbid MINI-Plus) psychotic disorder (including schizoaffective disorder) or with a dependency on alcohol or drugs within 12 months prior to hospitalization were excluded. Moreover, patients with insufficient mastery of the Dutch language were not deemed eligible. As DTW analyses focus on changes in severity scores between two adjacent assessments, these are more robust when there are 4 or more assessments. Therefore, we excluded patients with less than four BDI-II measurements, resulting in 166 inpatients in the current study (60.1%; n=1 with 0 measurements, n=4 with 1 measurement, n=36 with 2 measurements, and n=42 with 3 measurements). Patients were hospitalized at a depression ward and received treatment according to evidence-based guidelines (www.trimbos.nl, www.nvvp.net) (pharmacotherapy, (group)psychotherapy or a combination of both). Psychotropics at baseline were coded into five dichotomous variables: antidepressants, mood stabilizers, antipsychotics, benzodiazepines, and stimulants.

Beck Depression Inventory-II (BDI-II)

The administration of the BDI-II was part of a ROM test battery assessing mood, quality of life and overall mental well-being [17]. Routine Outcome Monitoring was performed at admission and every 2 weeks thereafter until discharge from the hospital. The BDI-II data presented in this article represent the data acquisition from the period April 2015 through February 2018.

The BDI-II [19] (Dutch translation by van der Does) [21] is a widely used and validated questionnaire for screening depression. It is a self-report inventory consisting of 21 items measuring affective, cognitive, and somatic depressive symptom domains. Each item is rated on a 4-point Likert scale, ranging from 0 to 3 and with higher scores indicating a higher severity of depressive symptomatology. The internal consistency of the BDI-II is adequate with a reported Cronbach's alpha of 0.91 [22,23]. In our sample, the omega values were 0.84, 0.90, and 0.93, at baseline and 2 and 4 weeks admission, respectively.

Statistical analyses

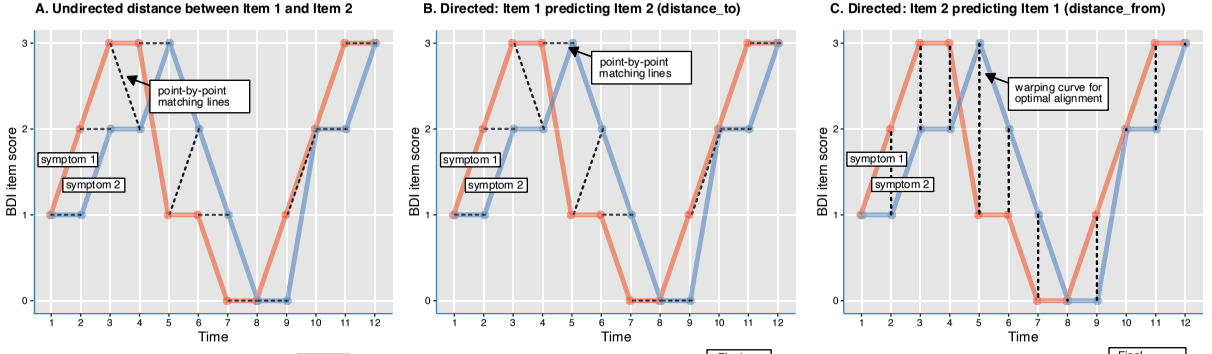
Sociodemographic and clinical variables are summarized in Table 1 as means with standard deviations (SD) or number (percentages), as appropriate.

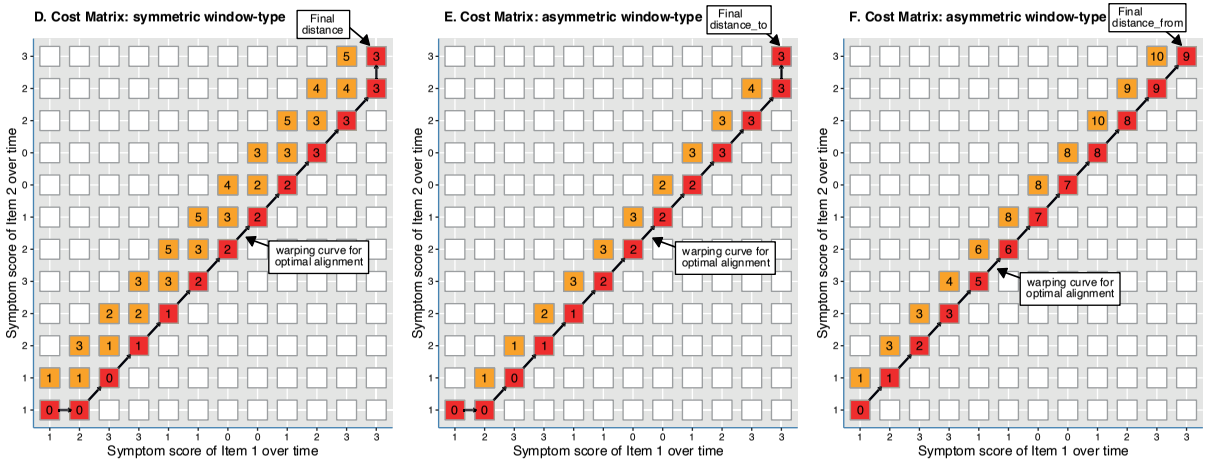
DTW analysis was used to identify BDI-II symptom dynamics in time [17,18]. Four steps were taken in the analyses. First, undirected analyses were done on the nomothetic group-level (based on aggregation of individual-level findings). Second, directed analyses were done in each of the inpatients separately yielding 166 asymmetric (distance) matrices. The results from three of these inpatients were presented as idiographic examples. Third, directed nomothetic analyses were done through bootstrapping, yielding a directed symptom network of the significant edges.

All item scores were group-level standardized before the analyses (except for Figure 1, for explanatory purposes only), in order to let the results be solely based on the change profiles over time.

Undirected DTW analyses

For the undirected analyses, DTW was used as an algorithm to calculate the "distance" between each pair of items, resulting in a 21 by 21 symptom distance matrix for each individual. The time window was symmetric and set to 1, meaning that changes between t-1 and t+1 were taken into account. A programming approach with elastic stretching and compressing was used in order for the two vectors of panel data to become optimally aligned through a warping path [24]. Changes in item 2 that occurred one time interval before or after that of item 1 could be stretched so that they overlapped in the DTW algorithm (Figure 1A). This resulted in 166 symmetric distance matrices, one per person.

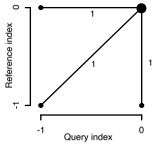




F. Step pattern: "symmetric1"

Recursion:

g[i , j] = min (
g [i-1 , j-1] + d [i , j] ,
g [i , j-1] + d [i , j] ,
g [i-1 , j] + d [i , j])



G. Calculation of the directed distances

$$\text{Distance}_{\text{item1 to item 2}} = \frac{\text{distance\_from - distance\_to}}{\text{mean(distance\_from, distance\_to)}} = \frac{9-3}{(9+3)/2} = \frac{6}{6} = \mathbf{1.0}$$

$$\text{Distance}_{\text{item 2 to item 1}} = \frac{\text{distance\_to - distance\_from}}{\text{mean(distance\_from, distance\_to)}} = \frac{3-9}{(9+3)/2} = \frac{-6}{6} = \mathbf{-1.0}$$

Figure 1. Explanation of undirected and directed DTW analysis of one symptom pair
Explanation of the directed dynamic time warp (DTW) analysis, an algorithm for measuring similarity between two time series. DTW is based on the concept of a warping curve, that stretches the two given time series so that they will overlap. Assume two BDI symptom scores over time from one individual with 12 time points (t). First, DTW creates a local cost matrix (LCM) with t x t dimensions (here, 12 x 12, panels D, E, and F). Second, DTW finds the path that minimizes the alignment between the two scores by iteratively stepping through the LCM, starting at the lower left corner (i.e., LCM[1, 1]) and finishing at the upper right corner (i.e., LCM[12, 12]), while aggregating the total distance (i.e., 'cost'). At each step, the algorithm takes the step in the direction in which the cost increases the least under the chosen constraint (i.e., Sakoe-Chiba window of size one=and "symmetric1" step pattern). In panel A, the undirected distance between the (unstandardized) scores of these individual symptoms s1 and s2 are assessed (with a symmetric time window). In panels B and C, the stretching is only allowed in one direction, after the current assessment, to yield directed distances. In the yellow boxes, the calculation of the directed distance is explained, which is the distance difference divided by the average distance of the symptom pair, yielding 1.0 for symptom 1 to symptom 2, and -1.0 for symptom 2 to symptom 1. Thus, we can conclude that the changes of the red symptom 1 scores tended to precede the changes of the blue item 2 scores.

Dissimilar scores at the start and end of each panel data could have a disproportional effect on the total distance because these cannot be dynamically aligned. Therefore, we used interpolation of 5 values between each time point before calculating the distance, which subsequently reduced the disrupting effect of starting and endpoints mismatches but did not affect the relative rank order of each of the 21*21 distances within each individual (see Supplementary R script). Moreover, a penalty of 2 was added to each of the original symptom pairs that both scored zero at the same instance, to reduce the tendency of item scores that remain constant throughout follow-up to cluster strongly together.

The 166 distance matrices were subsequently analyzed on the group level for the undirected analysis through a Distatis analysis (Figure 2) [25]. The aim of this method is to find the stable part in the similarities of the symptom dynamics between participants. It is a three-way principal component analysis, of the array of 166 distance matrices. The Distatis analysis yielded compromise factors, of which the first three scores best describe the similarity structure of the 166 distance matrices. The first against the second and the first against the third compromise factors were plotted into the x-y planes, and the three compromise factors were also plotted in a supplemental three-dimensional interactive plot (using the 'plotly' r package). To estimate the optimal number of dimensions, elbow and silhouette plots were being used to yield the optimal number of dimensions. The elbow can be observed as a sharp change in the slopes of adjacent line segments, which location might indicate a good number of dimensions to retain. The silhouette method calculates the average distance of each item to all the items in the same dimensions as well as the average distance to all the items in the nearest cluster, with a plot of the average scores over all items against different number of dimensions. The number of dimensions yielding the highest average silhouette score is the best number of dimensions [26]. To assess

which of the 21 items clustered together in dimensions, a hierarchical cluster analysis was applied according to 'Ward.D2' clustering methods, which was visualized in a dendrogram.

To assess whether our analytical approach yielded reliable results, we performed a random split of the data and repeated the analyses in both subsets. Through Procrustes analysis the two networks were brought into a similar space in which statistically meaningless differences were removed without changing the fit. This helped us to determine whether this resulted in similar findings or discrepant results, which may signal an inherently faulty method (Figure 3). The congruence coefficients (with the 2.5th and 97.5th percentiles) were estimated, through bootstrapping of 200 random splits of the 166 participants. A value below 0.85 is indicates poor similarity, a value in the range of 0.85 to 0.94 indicates fair similarity, and a value of 0.95 can be considered as being equal [27].

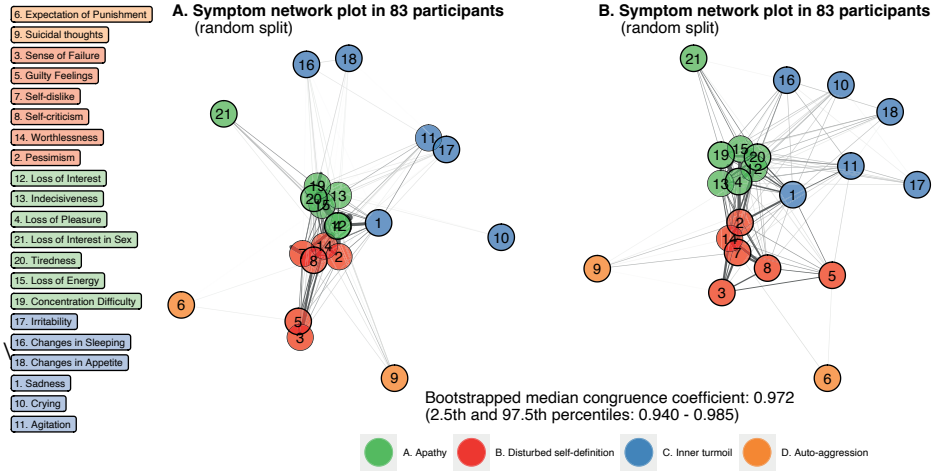


Figure 3. Undirected network plots of two random subsamples of the 166 inpatients
An automated split yielded subsamples of 83 and 83 patients each, in which we conducted separate dynamic time warp analyses. Network configurations are shown for both subsamples. The median congruence coefficient was very high at 0.972 (with the 2.5th and 97.5th percentiles of 0.940 and 0.985) when we bootstrapped the random split procedure 200 times, indicating stability of the 4 nomothetic symptom dimensions across participants.

Directed DTW analyses

For the directed analyses, we used the same DTW algorithm as before, with one crucial difference (see also Figure 1B and 1C): the window type using the Sakoe-Chiba band was specified as being asymmetric in order for the dynamic alignment to be only possible in one direction of time. The positive relative difference between those two distances will be the final distance (i.e., distance from item 1 to item 2). The higher this outcome is, the stronger the temporal effect from item 1 to item 2. For each of the 166 patients, a directed distance matrix was calculated. These analyses yielded directed symptom networks and outstrength and instrength centrality values, which are shown for 3 sample patients in Figure 4.

From each of the 166 individual directed network plots, the significant edges were assessed. Using bootstrapping with 9999 resamples with replacement, the mean and 95% confidence intervals (i.e., 2.5th and 97.5th percentiles) were assessed for each of the 21 x 21 = 441 distances. The directed network was plotted showing the statistically significant edges. The directed edges with arrow tips at one end of each edge denote the temporal direction of the effect and can be interpreted as Granger-causal (e.g., item 1 is statistically related to item 2 and changes in item 1 temporally precede changes in item 2). Next, two metrics of node centrality were derived, the in- and out strength centrality values. Outstrength centrality refers to the number of edges that depart from a specific node. In our DTW analysis a symptom with a high score is one for which its changes are followed by many other symptom changes. Instrength centrality, on the other hand, refers to the number of incoming edges of a specific node. Using bootstrapping, the confidence intervals (i.e., 2.5th and 97.5th percentiles) for each of in- and out strength value for each the items were estimated.

Software packages and codes

The "psych" (version 2.0.12), "dtw" (version 1.22-3), "parallelDist" (version 0.2.4), "plotly" (version 4.10.0), "boot" (version 1.3-27), and "qgraph" (version 1.6.9) packages for the R statistical software were used (R version 4.1.1; R Foundation for Statistical Computing, Vienna, Austria, 2016. URL: https://www.R-project.org/). Code can be found in supplemental materials.

Results

Patient characteristics

We included all patients from the Hercules Routine Outcome Monitoring (ROM) study project [17] with a minimum of four consecutive BDI-II measurements. The 110 excluded inpatients did not differ from the 166 included inpatients for age (mean 52.3 versus 50; p=0.28) and gender (p=0.33). Demographic and clinical characteristics of the included patients are shown in Table 1, as well as the use of psychotropic medication. Patients were on average 50.1 years (± 14.8), and 104 were women (62.7%). A bipolar disorder was diagnosed in 28 patients (16.9%). The mean duration of illness was 11.2 ± 15 years. For 36 patients (21.7%), the current episode was the first depressive episode. The average baseline BDI-II score was 34.3 (± 9.1). Of the 166 patients, 134 (71.1%) showed treatment response (defined by a 50% reduction on the HRSD-17 score) and 90 (54.2%) remission (i.e., HRSD-17 score ≤ 7) at the end of admission. The median duration of hospitalization was 14.5 weeks.

Table 1: Demographic and clinical characteristics of the sample

Variable	Mean (SD) / No (%)
Demographic characteristics	
Female (%)	104 (62.7)
Bipolar disorder	28 (16.9)
Age in years, mean (SD)	50.1 ± 14.8
Education:[a,b] (n, %)	
- Lower	30 (18.3)
- Intermediate	71 (43.3)
- Higher	63 (38.4)
Work status: (n, %)	
- Unemployed	90 (54.2)
- Employed	68 (41.0)
- Other (student/voluntary service)	8 (4.8)
Marital status: (n, %)	
- Married	70 (42.2)
- Divorced/widowed	45 (27.1)
- Never married	51 (30.7)
Living situation:[c] (n, %)	
- Living alone	51 (30.7)
- Living with partner	60 (36.1)
- Living with family	55 (33.1)

Clinical characteristics	
Index (first) depressive episode (n, %)	36 (21.7)
History of 4 or more depressive episodes (n, %)	44 (26.5)
Lifetime substance abuse/dependency:[d]	
- Alcohol	18 (10.8)
- Drugs (THC, hard drugs, benzodiazepines)	4 (2.4)
Melancholic features	102 (61.5)
ROM baseline total scores:	
- HRSD-17	20.3 ± 4.8
- BDI-II	34.3 ± 9.1
Baseline medication use:	
- Antidepressants	134 (80.7)
- Antipsychotics	74 (54.6)
- Mood stabilizers	24 (14.5)
Responders (%)	118 (71.1)
Remitters (%)	90 (54.2)

Data are mean (SD) or No (%), when appropriate. ROM: routine outcome monitoring, HRSD-17: Hamilton Depression Rating Scale, BDI-II: Beck Depression Inventory.
[a]Lower education: general basic education only; intermediate education: middle vocational education; higher education: higher vocational education or university. [b]Two missing values for education.
[c]Living alone includes living in a home for elderly and the convent
[d]Investigated using the MINI modules on substance abuse

Undirected DTW analyses

Group-level undirected DTW analyses were conducted on the change profiles of the BDI-II symptoms. Figure 2 shows these results for the 166 depressed inpatients. Distatis analysis yielded 21 compromise factors, of which the first three explained 9.5%, 8.6%, and 7.1% of the variance, respectively (Figure 2A and 2B). Based on the elbow and silhouette methods, the most parsimonious number of dimensions was four (Figure 2C). The hierarchical cluster analyses yielded four symptom dimensions (Figure 2D) which we termed 'A. Apathy' (7 items), 'B. Disturbed self-definition' (6 items), 'C. Inner turmoil' (6 items), and 'D. Auto-aggression (2 items). The three compromise factors are also depicted in a three-dimensional plot (Supplementary Figure 1).

The shortest distances (representing the most similar symptom changes within patients over time) were between the item pairs "15. Loss of Energy" and "19. Concentration Difficulty", and between "2. Pessimism" and "14. Worthlessness". The longest distances (i.e., most dissimilar) were between the item pairs "4. Loss of Pleasure" and "6. Expectation of Punishment".

To obtain an idea of the accuracy of results, we obtained a median congruence value of 0.972 (95% CI: 0.940 - 0.985) through 200 bootstrap analyses in two datasets of n=83 each (Figure 2). As the median congruence value is larger than 0.95, the network can be regarded as very stable.

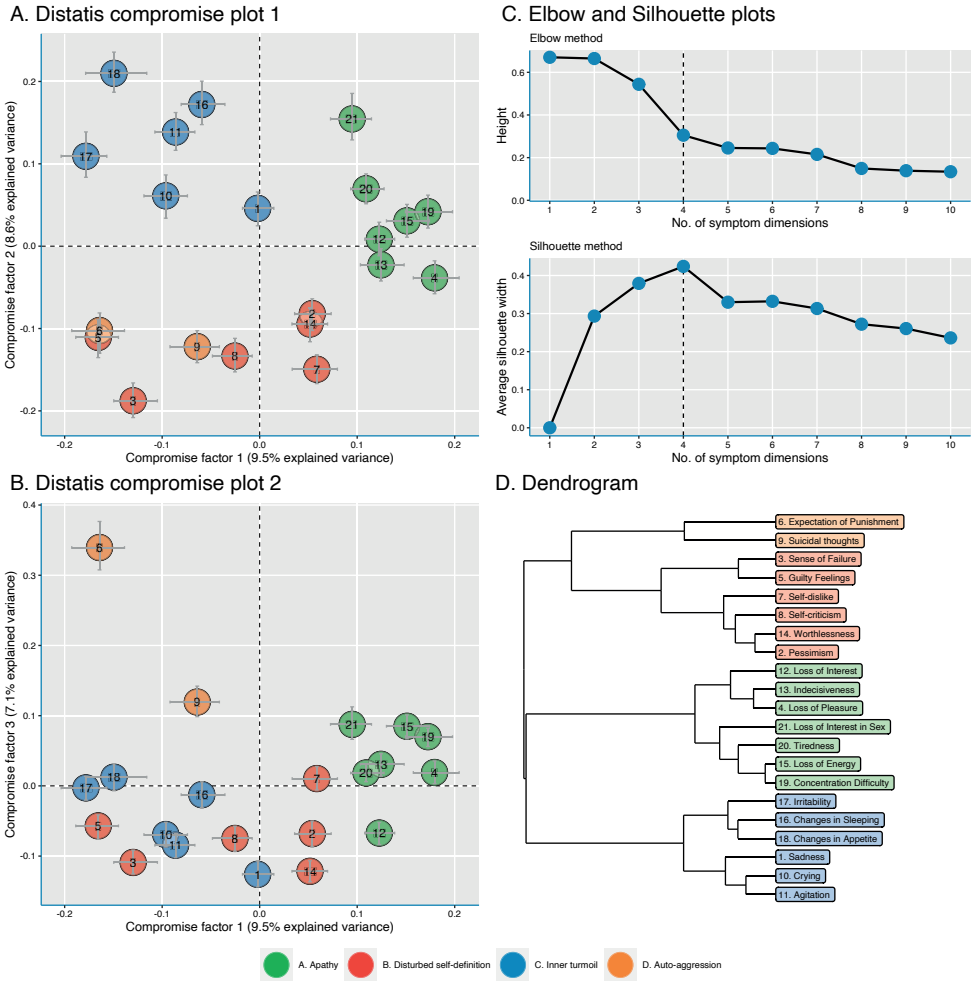


Figure 2. Nomothetic undirected DTW analysis of BDI-II symptoms

Panels A and B show the compromise plots based on the Distatis analysis (three-way principal component analysis of the 166 distance matrices) that shows the position of the 21 BDI-II items in the compromise space using the first 2 compromise factors (panel A) and the first and the third compromise factor (panel B). The grey horizontal and vertical error bars represent the 95% confidence intervals, assessed through estimated through bootstrapping with 500 resamples. Panel C shows the scree plot based on the eigenvalues in a downward curve based on three compromise factors. The number of dimensions was determined using the elbow method (i.e., the point where the slope of the curve is levelling off is after 4 dimensions) and was also determined using silhouette plot (i.e., 4 dimensions yielded the highest average silhouette score). Panel D shows the dendrogram based on the hierarchical clustering procedure based on three compromise factors.

Directed DTW analyses

Figures 1B and 1C explain the calculation of the directed DTW analysis of one BDI item pair, which is contrasted with that of the undirected DTW analysis (Figure 1A). The directed distance matrices were analyzed for each of the 166 patients, separately. The idiographic analyses of three of these are presented in Figure 4. These show that patients differ according to their symptoms with high out- or instrength: for example in patient 91 the 'Apathy' symptoms show highest outstrength, whereas in patient 129 the symptoms belonging to dimension 'Disturbed self definition' show the highest outstrength.

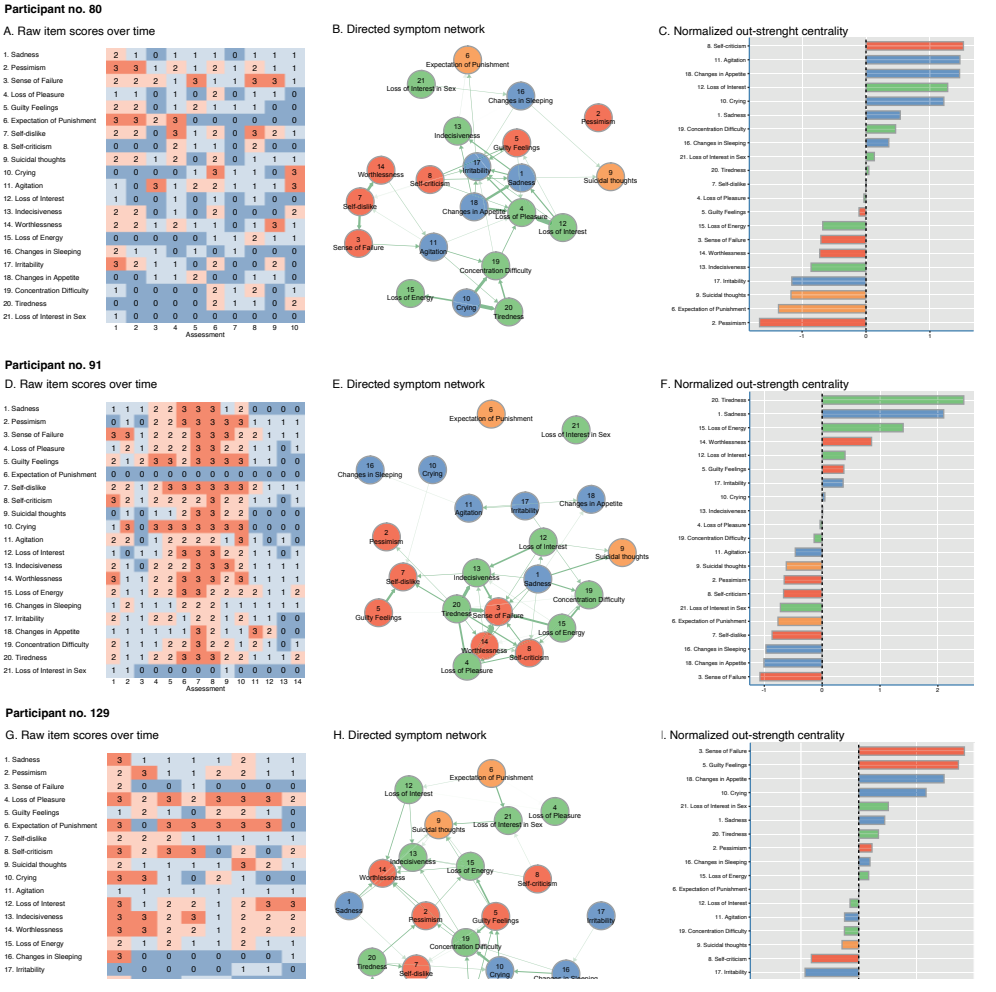


Figure 4. Idiographic directed DTW analyses of 3 depressed inpatients

Idiographic directed DTW analyses based on all distance matrices based on change profiles of BDI-II symptom scores from 3 of the 166 depressed inpatients.

A. The crude scores of each of 21 BDI-II items over time.

B. The idiographic directed network plot of the 21 BDI-II symptoms.

C. The normalized outstrength centrality of the 21 BDI-II symptoms.

Figure 5A shows the nomothetic directed network. The symptoms with the significantly strongest out-strength values were 'Worthlessness' and 'Pessimism'. Figure 5B depicts the in- and outstrength centrality of this network. The symptoms 'Worthlessness', 'Pessimism', 'Loss of interest', 'Tiredness', and 'Sadness' showed outstrength values significantly larger than the average outstrength value. This means that when patients report changes in worthlessness and pessimism, it is probable that they will also report changes in several other symptoms at the next assessment. Symptoms belonging to the 'Apathy' dimension had relatively large instrength centrality values. The items 'Loss of interest' and 'Self-criticism' had significantly higher than average instrength centrality values, meaning that their fluctuations mostly followed upon similar fluctuations in other symptoms.

Discussion

This study used DTW to investigate the dynamics of depressive symptoms in severely depressed patients. It was based on the premise that BDI-II symptoms do not co-occur in synchronicity, and that some are more likely to covary than others, in part because of complex interactions among symptoms [7,15]. We showed that not only an undirected network, but also a directed network could be constructed using panel data with only a sparse number of datapoints per patient. We found high outstrength centrality values of 'Worthlessness', 'Pessimism', 'Loss of interest', 'Tiredness', and 'Sadness'—these symptoms hence have a higher likelihood of predicting similar fluctuations in other symptoms. Symptoms of the 'Apathy' dimension had relatively highest instrength centrality values and were more likely to be susceptible to change upon fluctuations in other symptoms. Thus, DTW shows promise to extend the arsenal of analytic methods to analyze individual longitudinal patient data, which data could subsequently be aggregated into group-level analyses. Thereby, it could help to deepen our understanding through a complex dynamic systems lens [7,28].

Group-level undirected DTW analyses revealed four symptom dimensions with similar dynamics over time: 'A. Apathy', 'B. Disturbed Self-definition', 'C. Inner turmoil', and 'D. Auto-aggression. Symptoms within the same cluster tended to change in synchronicity, more so than with symptoms from the other three clusters. How do our results compare to the cross-sectional factor analytic studies of the BDI-II? While the literature is inconsistent, several studies have pointed towards the existence of a Somatic-Affective and Cognitive factor [29]. Our 'Disturbed Self-definition' cluster showed a clear overlap with the Cognitive (or melancholic) factor (only suicidality was not found in our cluster). This was in agreement with findings from the study of Bringmann et al. [30] who found 2 clusters: 'Cognitive' cluster and 'Somatic-Affective' cluster.

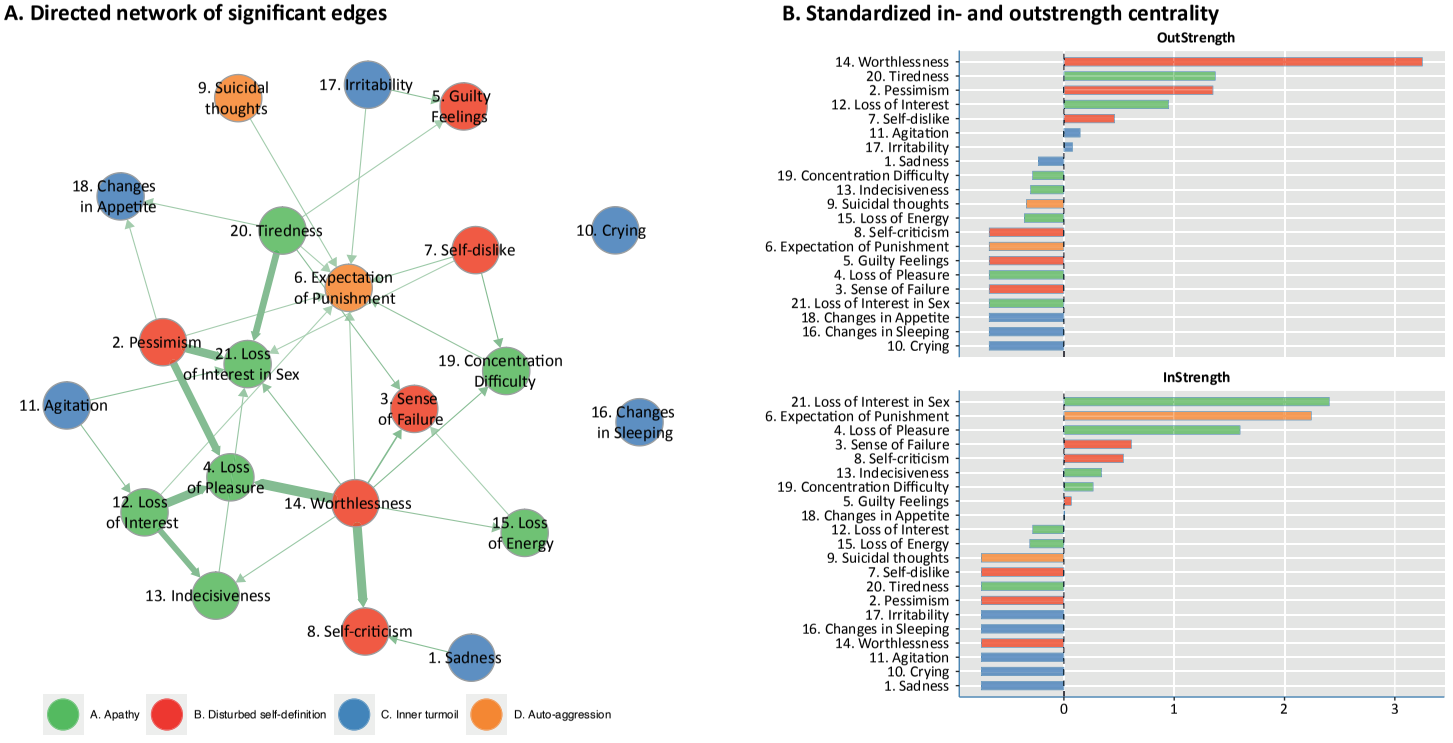


Figure 5. Nomothetic directed DTW analyses

The directed network plot is based on the 166 directed distance matrices of BDI-II symptom scores from 166 depressed inpatients. The mean and 95% confidence intervals (i.e., 2.5th and 97.5th percentiles) were estimated through bootstrapping with 1000 resamples. Only the statistically significant edges are shown. The edge thickness signifies the magnitude of the temporal associations. The standardized in- and out-outstrength values were based on the statistically significant edges in the directed symptom network.

Our group-level directed DTW analyses showed that the symptoms 'Worthlessness', 'Pessimism', 'Loss of interest', 'Tiredness', and 'Sadness' had large outstrength relative to other items. From a complex dynamic systems perspective, this could indicate that these items have a positive feedback effect on other symptoms. Bringmann et al. [15] found 'Loss of Pleasure' to be the symptom with the largest outstrength, and also low in-strength, but they did not find 'Sadness' to be a highly outstrength symptom. Symptoms with a high outstrength could have interesting clinical implications as these symptoms could play a critical role in transition towards the onset or resolution networks of depressed symptoms. Not many previous studies have used longitudinal designs, but in a cross-sectional analysis on BDI-II centrality in depressed patients, 'Sadness' and 'Pessimism' were also the most central BDI-II symptoms. Interestingly, they further showed that 'Pessimism' was the most important predictive item for future outcome. Importantly, it must be acknowledged that the DTW method for assessing symptoms dynamics in psychopathology is still in its infancy and the added value compared to other statistical analysis techniques must be further evaluated in follow-up studies.

The methods of a directed DTW analysis may increase our insight in the symptom dynamics in depression in light of complex dynamic systems theory [31]. We need a better understanding of the dynamic processes that lie at the basis of changes in mood states (e.g., critical transitions or phase transitions). In order to do this, we need to combine individual-level and group-level approaches. The traditional group-level approaches focus merely on the mean changes in depression symptom severity, wherein important information on the dynamic symptom interactions is lost.

A main strength of this study was the longitudinal design and the novel analytical DTW approach, which provides a means to study symptom dynamics without requiring high intensity measurements. Furthermore, the DTW output can be visualized which can make idiographic analyses comprehensible both for the individual patient as well as the clinician. Advantages of the DTW approach are, in contrast to the VAR-1 lagged models [15]: it is a non-linear analytic technique; the stationarity criterion [10] does not have to be met; DTW needs less assessments per participant. Advantages of other data analytic methods such as VAR-1 lagged models or GIMME [32] is that they analyze multivariate relationships but they need many more assessments per participants which is not always feasible in clinically depressed patients.

Regarding the limitations, first, assessments were taken at an interval of 2 weeks apart, which precluded the detection of faster-moving processes. Second, there were only on average 7 assessments per patient, whereas the individual outcomes will be more robust when there would have been more assessments. Third, we set rather arbitrary thresholds in the two network plots showing only the strongest edges. Moreover, only

estimated positive temporal associations (i.e., positive feedback loops), whereas it may also be of importance to get insight into whether increases in one symptom are followed by decreases in some other symptom (i.e., negative feedback loops). Fourth, inpatients received treatment that likely affected the networks, which calls for adaptions of DTW that can estimate time-varying networks. Finally, it should be noted though that some symptoms may be functionally indistinguishable from one another (e.g., 'Self-criticism' and 'Self-dislike') [33].

From a clinical perspective, the idiographic networks may provide an increased understanding of the complex interactions between symptoms. These findings could be visualized, which may be of clinical value to patients and clinicians untrained in advanced statistics. Further research is needed to study whether targeting symptoms with high outstrengths is of benefit in clinical practice. However, selectively deactivating a symptom is more easily said than done in psychiatry [10]. Further studies should also aim to integrate transdiagnostic symptoms and other variables (e.g., environmental factors) to gain insight in the potential temporal relations of psychopathology.

Supplemental Material

Supplementary Figure 1. Three-dimensional compromise plot of 21 BDI-II symptoms
https://osf.io/gzd7w/

A Distatis analysis was done on the 166 distance matrices based on change profiles of BDI-II symptom scores from 166 depressed inpatients yielding three compromise factors (i.e., principal components), which are depicted in a 3D interactive scatter plot. The items are represented as points such that the distances best reflect the similarities between the change profiles of the items (Abdi et al. 2012).

Supplementary R script
https://osf.io/ahqfz/

References

1. Borsboom D, Cramer AOJ. Network Analysis: An Integrative Approach to the Structure of Psychopathology. *Annu Rev Clin Psychol*. 2013;9(1):91-121. doi:10.1146/annurev-clinpsy-050212-185608
2. Fried EI, Nesse RM, Zivin K, Guille C, Sen S. Depression is more than the sum score of its parts: individual DSM symptoms have different risk factors. *Psychol Med*. 2014;44(10):2067-2076. doi:10.1017/S0033291713002900
3. Fried EI, Nesse RM. The impact of individual depressive symptoms on impairment of psychosocial functioning. *PLoS One*. 2014;9(2). doi:10.1371/journal.pone.0090311
4. Kendler KS, Aggen SH, Neale MC. Evidence for multiple genetic factors underlying DSM-IV criteria for major depression. *JAMA Psychiatry*. 2013;70(6):599-607. doi:10.1001/jamapsychiatry.2013.751
5. Myung W, Song J, Lim SW, et al. Genetic association study of individual symptoms in depression. *Psychiatry Res*. 2012;198(3):400-406. doi:10.1016/j.psychres.2011.12.037
6. Borsboom D, Deserno MK, Rhemtulla M, et al. Network analysis of multivariate data in psychological science. *Nat Rev Methods Prim*. 2021;1(1). doi:10.1038/s43586-021-00055-w
7. Cramer AOJ, Van Borkulo CD, Giltay EJ, et al. Major depression as a complex dynamic system. *PLoS One*. 2016;11(12). doi:10.1371/journal.pone.0167490
8. Fried EI, Robinaugh DJ. Systems all the way down: Embracing complexity in mental health research. *BMC Med*. 2020;18(1):4-7. doi:10.1186/s12916-020-01668-w
9. Robinaugh DJ, Hoekstra RHA, Toner ER, Borsboom D. The network approach to psychopathology: A review of the literature 2008-2018 and an agenda for future research. *Psychol Med*. 2020;50(3):353-366. doi:10.1017/S0033291719003404
10. Mcnally RJ. Network Analysis of Psychopathology: Controversies and Challenges. *Annu Rev Clin Psychol*. 2021;17:31-53. doi:10.1146/annurev-clinpsy-081219-092850
11. Nelson B, McGorry PD, Wichers M, Wigman JTW, Hartmann JA. Moving from static to dynamic models of the onset of mental disorder a review. *JAMA Psychiatry*. 2017;74(5):528-534. doi:10.1001/jamapsychiatry.2017.0001
12. Van De Leemput IA, Wichers M, Cramer AOJ, et al. Critical slowing down as early warning for the onset and termination of depression. *Proc Natl Acad Sci U S A*. 2014;111(1):87-92. doi:10.1073/pnas.1312114110
13. Fisher AJ, Medaglia JD, Jeronimus BF. Lack of group-to-individual generalizability is a threat to human subjects research. *Proc Natl Acad Sci U S A*. 2018;115(27):E6106-E6115. doi:10.1073/PNAS.1711978115
14. Heino M, Knittle K, Noone C, Hasselman F, Hankonen N. Studying Behaviour Change Mechanisms under Complexity. *Behav Sci (Basel, Switzerland)*. 2021;11(5). doi:10.3390/BS11050077

15. Bringmann LF, Vissers N, Wichers M, et al. A Network Approach to Psychopathology: New Insights into Clinical Longitudinal Data. *PLoS One*. 2013;8(4):e60188. doi:10.1371/JOURNAL.PONE.0060188
16. Hekler E, Klasnja P, Chevance G, Golaszewski N, Lewis D, Sim I. Why we need a small data paradigm. *BMC Med*. 2019;17(1). doi:10.1186/S12916-019-1366-X
17. Hebbrecht K, Stuivenga M, Birkenhäger T, et al. Understanding personalized dynamics to inform precision medicine: a dynamic time warp analysis of 255 depressed inpatients. *BMC Med*. 2020;18(1):1-15. doi:10.1186/s12916-020-01867-5
18. Booij M, van Noorden M, van Vliet I, et al. Dynamic time warp analysis of individual symptom trajectories in depressed patients treated with electroconvulsive therapy. *J Affect Disord*. 2021;293:435-443. doi:10.1016/J.JAD.2021.06.068
19. Beck AT, Steer RA BO. *Beck Depression Inventory Manual, 2nd Ed*. Psychological Corporation, San Antonio, TX; 1996.
20. Fried EI, Flake JK, Robinaugh DJ. Revisiting the theoretical and methodological foundations of depression measurement. *Nat Rev Psychol*. 2022. doi:10.1038/s44159-022-00050-2
21. *Van Der Does AJW (2002). Manual of the Dutch Version of the Beck Depression Inventory (BDI-II-NL). Harcourt Publishers: Amsterdam.*
22. Schotte CKW, Maes M, Cluydts R, De Doncker D, Cosyns P. Construct validity of the Beck Depression Inventory in a depressive population. *J Affect Disord*. 1997;46(2):115-125. doi:10.1016/S0165-0327(97)00094-3
23. Uher R, Farmer A, Maier W, et al. Measuring depression: comparison and integration of three scales in the GENDEP study. *Psychol Med*. 2008;38(2):289-300. doi:10.1017/S0033291707001730
24. Giorgino T. Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package. *J stat*. 2009;31(7):1-24.
25. Abdi, H; Valent, D; Chollet, S; Chrea C. Analyzing Assessors and Products in Sorting Tasks: DISTATIS, Theory and Applications. . 2007;18:1-16. *Food Qual Pref*. 2007;181:1-16.
26. Kingrani SK, Levene M, Zhang D. Estimating the number of clusters using diversity. *Artif Intell Res*. 2017;7(1):15. doi:10.5430/AIR.V7N1P15
27. Lorenzo-Seva U, ten Berge JMF. Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology*. 2006;2(2):57-64. doi:10.1027/1614-2241.2.2.57
28. Olthof M, Hasselman F, Lichtwarck-Aschoff A. Complexity in psychological self-ratings: implications for research and practice. *BMC Med*. 2020;18(1). doi:10.1186/S12916-020-01727-2
29. Vanheule S, Desmet M, Groenvynck H, Rosseel Y, Fontaine J. The factor structure of the Beck Depression Inventory-II: an evaluation. *Assessment*. 2008;15(2):177-187. doi:10.1177/1073191107311261
30. Bringmann LF, Lemmens LHJM, Huibers MJH, Borsboom D, Tuerlinckx F. Revealing the dynamic network structure of the Beck Depression Inventory-II. *Psychol Med*. 2015;45(4):747-757. doi:10.1017/S0033291714001809
31. Ehlers CL. Chaos and complexity. Can it help us to understand mood and behavior? *Arch Gen Psychiatry*. 1995;52(11):960-964. doi:10.1001/ARCHPSYC.1995.03950230074010

32. Beltz AM, Wright AGC, Sprague BN, Molenaar PCM. Bridging the Nomothetic and Idiographic Approaches to the Analysis of Clinical Data. *Assessment*. 2016;23(4):447-458. doi:10.1177/1073191116648209
33. Fried EI, Cramer AOJ. Moving Forward: Challenges and Directions for Psychopathological Network Theory and Methodology. *Perspect Psychol Sci*. 2017;12(6):999-1020. doi:10.1177/1745691617705892

CHAPTER 7

Cognitive change after electroconvulsive therapy in mood disorders measured with the Montreal Cognitive Assessment

Hebbrecht K, Giltay EJ, Birkenhäger TK, Sabbe B, Verwijk E, Obbels J, Roelant E, Schrijvers D, Van Diermen L.

Acta Psychiatr Scand. 2020 Nov;142(5):413-422.

Abstract

Objective: The Montreal Cognitive Assessment (MoCA) is a sensitive and clinically practical test but its usefulness in measuring long-term cognitive effects of ECT is unclear. Using the MoCA, we investigated short- and long-term global cognitive change in ECT-treated patients with a Major Depressive Episode (MDE).

Method: We included 65 consecutive ECT-treated patients with MDE, in whom global cognitive functioning was assessed at baseline (T0); during ECT (before the third session; T1); and 1 week (T2), 3 months (T3), and 6 months (T4) after completion of the index course. Changes in MoCA (sub)scores were analyzed using linear mixed models and reliable change indices were computed to investigate individual changes in MoCA total scores.

Results: There was a significant effect of time on MoCA scores ($F(4, 230.5) = 4.14$, $P = 0.003$), with an improvement in global cognitive functioning from T3 compared to T1 and T2. At the individual level, 26% (n = 17) of patients showed a significantly worse cognitive functioning at T2 and 12% (n = 8) an improved cognitive functioning compared to T0. For T4, these percentages ameliorated to 8% and 18% respectively.

Conclusion: No persistent global cognitive impairment induced by ECT was found at the group level using the MoCA. At the individual level, however, there was clear heterogeneity in the effects of ECT on cognitive functioning. The MoCA is a suitable tool to monitor short- and long-term global cognitive functioning in ECT-treated patients with MDE but in younger patients, potential ceiling effects must be taken into account.

Introduction

Electroconvulsive therapy (ECT) is an effective [1,2] and safe [3] biological treatment for patients experiencing a severe major depressive episode (MDE). However, there is concern about its neurocognitive side-effects [4,5]. Immediately following treatment, some patients experience post-ECT disorientation, which usually resolves within 1 h [6]. Subacute cognitive side-effects include anterograde amnesia for recently learned information and retrograde amnesia for previously learned information. The ECT-related brain disruption causing anterograde amnesia typically normalizes within two months after ECT completion [5,7]. Some patients experience persistent retrograde amnesia [8], but research regarding its occurrence and evolution is scarce [5,9].

Various guidelines recommend performing cognitive assessments before, during, and after a course of ECT [10], but there is no general consensus on timing or types of measurements to use [11–13]. The Mini-Mental State Examination (MMSE) [14] is the most widely used instrument to assess the impact of ECT on global cognitive functioning [5,13], but the MMSE lacks the sensitivity to detect more subtle forms of ECT-induced cognitive impairment [15].

The Montreal Cognitive Assessment (MoCA) may be a better alternative for two reasons. First, compared to the MMSE, the MoCA is a more sensitive cognitive screening instrument in various neurobehavioral (e.g., Parkinson's [16], Alzheimer's [17]) and psychiatric disorders (e.g., schizophrenia [18], depression [19]). Accurately identifying baseline (or pre-ECT) cognitive functioning is crucial to avoid the false attribution of cognitive impairment after ECT as ECT-induced. Second, the sensitivity of the MoCA regarding its ability to identify short-term ECT-induced cognitive impairment has been demonstrated in mixed psychiatric samples [4,20], in schizophrenia [21,22], and in MDE [15]. Compared to the MMSE, the MoCA provides a more valid and reliable evaluation of attention, visuo-executive functioning, and language fluency [23]—the last two domains are commonly affected by ECT in patients with MDE [5,24]. Furthermore, the MoCA assesses ECT-induced anterograde memory more extensively by using a delayed recall test that features five words instead of three [23].

Currently, there is insufficient evidence for an optimal MoCA cutoff for cognitive impairment. The initially suggested cutoff of 26/30 [23] has led to a high rate of false positives in several validation studies [25–28] and a recent meta-analysis has proposed the use of a cutoff of 23 [29]. The usefulness of the MoCA to detect cognitive change over time has been confirmed in previous studies [30,31].

When evaluating ECT-induced cognitive impairment, certain confounding, and effect-modifying factors should be taken into consideration [32]. Previous research has consistently shown that a bitemporal electrode placement is associated with greater cognitive side-

effects compared to a right unilateral placement [33,34]. Research has yet to confirm whether increasing the number of sessions, an older age, a lower educational level, the presence of pre-ECT cognitive impairment, psychotic symptoms, or a more severe depression can influence the amount of cognitive side-effects [32,35]. Vulnerability to ECT-induced cognitive impairment varies greatly on an individual level [32,36]. which has led to the recent call to investigate individual cognitive effects (aside from group means) of global cognitive functioning [37,38].

Aims of the study

This current study is the first that uses alternate versions of the Montreal Cognitive Assessment to investigate both short- (during Electroconvulsive therapy and directly after completion) and longterm (three and six months after completion) effects of ECT on global cognitive functioning as assessed by the MoCA in a sample of patients with major depressive episode. A second aim was to assess the individual variation of ECT-induced cognitive functioning using reliable change indices (RCI).

Material and methods

Study sample

We included patients with a major depressive episode (MDE) (Major Depressive Disorder or bipolar depression according to the DSM-IV-TR) who were 18 years and over. We recruited patients from the in- and outpatient departments of the University Psychiatric Hospital in Duffel (Belgium). Diagnoses were confirmed by using the MINI diagnostic interview version 6.0 [39]. To be included, patients had to score ≥17 on the Hamilton Depression Rating Scale-17 items (HDRS-17) [40]. We excluded patients with a history of substance abuse (<6 months prior) or who had been diagnosed with primary psychotic or schizoaffective disorders. Our study was part of the PROTECT study, which was designed to investigate ECT response predictors [41–43]. All patients provided written informed consent before the study procedures were performed. The study protocol complied with the Declaration of Helsinki and was approved by the local medical ethics committee. The study is registered in the online clinical database ClinicalTrials.gov (Identifier: NCT02562846).

Assessments

Cognitive assessments.

We used the MoCA to assess patients' global cognitive functioning in the week prior to ECT (T0), during ECT (before the third session, T1), within 1 week after completing the index course (T2), and at 3 (T3) and 6 (T4) months after T2. In order to prevent dropout, we tested patients who were unwilling or unable to travel to the hospital at home. The MoCA takes approximately 10 min to complete and assesses a wide range of cognitive domains,

which include visuo-executive functioning, naming, attention, language (repetition, fluency), abstraction, delayed recall, and orientation. The MoCA total score, obtained by the sum of scores on each completed task, ranges from 0 to 30—a one point educational correction (addition) is advised for individuals with ≤12 years of education [23]. Three different versions of the paper-and-pencil MoCA-Dutch version [44] were used and patients completed different versions of the MoCA at subsequent time point to minimize practice effects. Patients with a baseline MoCA score below 23 were designated as cognitively impaired [29].

Mood assessments.

We used the HDRS-17 to assess depression severity at each time point (T0–T4) as well as weekly during the ECT course. 'Response' was defined as a reduction of at least 50% on the HDRS-17 at T2 compared to baseline. 'Remission' was defined as scoring ≤7 on the HDRS-17 at T2.

Other assessments.

The CORE Assessment of Psychomotor Functioning was used to categorize patients into a melancholic and non-melancholic groups; a score of ≥8 indicated melancholic depression [45].

Treatment

Pharmacological treatment.

All participants continued their psychotropic medication regimen during the ECT course. When possible, drugs and doses were kept stable for 4 weeks before and during the ECT course.

Electroconvulsive therapy.

Patients were treated with ECT twice per week using a brief-pulse (0.5 ms), constant-current Thymatron IV system (Somatics LLC, Lake Bluff, Illinois). The electrodes were placed unilaterally over the right hemisphere, bifrontal, or bitemporal when a fast antidepressant effect was required or when patients did not respond to unilateral ECT. The ECT stimulus dose was determined by using the age method for right unilateral treatment and by using the halfage method for the bilateral electrode placement. Etomidate (0.15 mg/kg) was used for anesthetic induction, but propofol (1 mg/kg) was used when etomidate was not (well) tolerated. Ketamine (1–2 mg/kg) was used when there was no clinical response after 12 consecutive sessions. Succinylcholine (0.5 mg/kg) was used for muscle relaxation. ECT was terminated when patients were either in remission or showed no further clinical improvement during the last three ECT sessions.

Statistical analyses

Baseline demographic and clinical characteristics were compared between MoCA completers and non-completers (i.e. patients that missed one or more MoCA assessment) using chi-square tests for categorical variables and analysis of variance (ANOVA) for continuous variables.

To evaluate the effect of ECT on global cognitive functioning as assessed by the MoCA, a linear mixed model (LMM) was fitted with subject included as a random effect and time (time points T0 to T4) as a fixed effect. The same LMM was performed solely on the subsample that did not receive C-ECT after T2. We also fitted LMMs for each of the seven subscores of the MoCA (visuoexecutive functioning, naming, attention, language, abstraction, delayed recall, and orientation). As verbal fluency is only rated by a score of 0 or 1 in the second part of the language subitem (naming of words starting with a particular letter in 1 min; score 1 if a minimum of 11 words is possible), we added the maximum number of words that the patient could name as an additional outcome variable to the fluency task. When a significant change between time points emerged, we used Tukey's HSD corrections for multiple testing to conduct pairwise comparisons between all considered time points. Because we used eight different secondary outcome variables, we applied an additional Bonferroni–Holm correction to the P values of the fixed effect.

In order to verify the presence of possible ceiling effects of the MoCA, the number of measurements that reached the maximum score or a score of 28 or 27 over all time points was investigated.

As a supplementary analysis, we studied the potential effects of confounding and effect-modifying factors. We explored the effects of electrode placement (right unilateral, bilateral, mixed), age (age ≥ 65: yes/no), educational level (≤12 years of education: yes/no), presence of pre-ECT cognitive impairment (MoCA at T0 < 23: yes/no), psychotic symptoms (yes/no), severity of depression (at all time points), and the number of ECT sessions (at all time points) in seven separate LMMs. At first each covariate was added to the model as a main effect with inclusion of the interaction term between covariate and time point. If the interaction term was not significant the model was refitted without the interaction term and containing only the two main effects (time point and covariate).

We evaluated individual changes in cognitive functioning by using the reliable change index (RCI) [46], which indicates whether a change in score is significantly greater than expected based on test–retest reliability. We used a 90% confidence interval, meaning that we considered RCI values of 1.645 or higher statistically significantly. In this study, we used the method proposed by Jacobson and Truax—RCI = (posttest - pretest)/Standard Error (SE) [40].

We calculated the SE by using the variance in baseline scores of our sample and the reliability of the MoCA extracted from academic literature on MoCA's psychometric properties [23]. Baseline demographic and clinical characteristics were compared between patients that worsened, improved or remained stable on cognitive level at T2 in a supplementary analysis (using chisquare tests for categorical and ANOVA for continuous variables). All statistical analyses were performed in JMP Pro 14, with a significance level set at $P < 0.05$ (two tailed).

Results

Study population characteristics

A participant flow chart is presented in Figure 1. Demographic and clinical characteristics for the participants who completed the ECT treatment (n = 65) are shown in Table 1. There were no significant differences in demographic or clinical characteristic between patients who completed all MoCA assessments (n = 65) and patients who missed one or more MoCA assessment (n = 16). There were 50 responders (76.9%); remission was achieved in 40 patients (61.5%). Thirty-four (52.3%) patients received continuation ECT (CECT) during the 6-month time interval after completing the index course.

Table 1. Demographic and clinical characteristics of the study population

N	**65**
Age, years mean (± SD); range	58.4 (± 15.6); 21–85
Gender, female *n* (%)	50 (76.9)
Bipolar *n* (%)	12 (18.5)
Psychotic symptoms *n* (%) at baseline	31 (47.7)
Melancholia (CORE-defined) *n* (%)	41 (63.1)
Episode duration (Months)	
- Mean	13.3 (± 17.8)
- Median, range	6.0, 1–84
HDRS-17 mean (± SD) baseline (T0); range	24.9 (± 6.1); 17–38
Responders at T2 *n* (%)	50 (76.9)
Remitters at T2 *n* (%)	40 (61.5)
Cognitive impairment at baseline (MoCA < 23) *n* (%)	27 (41.5)
Number of ECT treatments in acute course, mean (± SD); range	11.2 (± 5.7); 2–27
Patients that received continuation ECT *n* (%)	34 (52.3)
Number of ECT continuation sessions (in *n* = 34 patients) mean (± SD); range	5.8 (± 8.5), 2–33
Electrode placement	
- Right unilateral *n* (%)	36 (55.4)
- Bilateral *n* (%)	13 (20)
○ Bifrontal *n* (%)	2 (3.1)
○ Bitemporal *n* (%)	11 (16.9)
- Mixed (Switch) *n* (%)	16 (24.6)
Mean stimulation dose (millicoulombs)	271.4

CORE Assessment of psychomotor functioning (CORE), Hamilton Depression Rating Scale (HDRS), Montreal Cognitive Assessment (MoCA), Electroconvulsive Treatment (ECT)

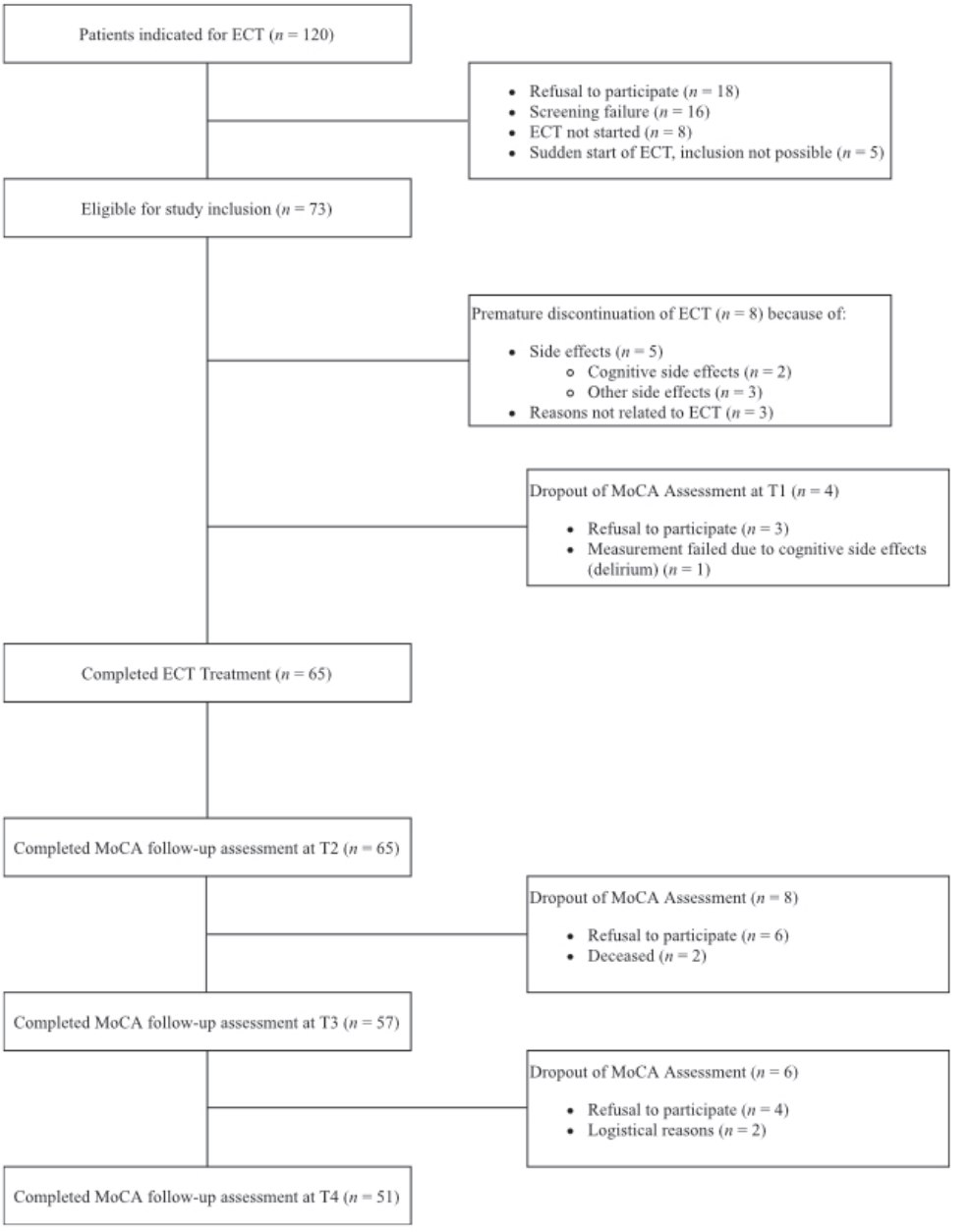


Figure. 1. Patient flow chart.

Evolution of global cognitive functioning on group level

An overview of MoCA total scores and subscores at the five time points can be found in Table 2. The change in these scores over time is shown in Figure 2. A significant change in MoCA total scores over the five time points was found (fixed effect time point, $F(4, 230.5) = 4.14$, $P = 0.003$). Tukey analyses revealed a significant improvement in MoCA total scores 3 months after ECT treatment (T3) compared to during (T1) and after treatment (T2), but MoCA total scores did not significantly improve compared to T0. These results were comparable when applying the analysis solely on the subsample that did not receive C-ECT after T2 ($F(4, 104.1) = 3.75$, $P = 0.007$).

There was a main effect of time, after Bonferroni–Holm correction, for the delayed recall task ($F(4, 231.9) = 3.68$; $P = 0.045$) and the number of words in the fluency task ($F(4, 230.7) = 3.89$; $P = 0.036$). Delayed recall performance improved significantly at T3 compared to T1 and T2. The score on the number-of-words additional outcome variable was significantly better at T3 compared to T2. The performance on the other six subitems (visuospatial and executive functioning, naming, attention, language, abstraction, and orientation) did not change significantly over the four time points.

The number of MoCA assessments that reached a score of more than 27 was only 39 out of 299 assessments (13%) which suggests that the ceiling effect of the MoCA was limited in our sample. 18 patients scored more than 27 on one or more time points (i.e. 39 measurements scored by 18 different patients) were significantly younger (44 ±14.2 (range: 21-67) than the patients that did not reach a minimum score of 27 on one or more time points (mean age 64±1.91; F-ratio: 27.74; $P < 0.001$). The MoCA subtasks Naming, Abstraction and Orientation did show ceiling effects.

Table 2. MoCA total mean scores and subscores before, during (before the third session), and after ECT as analyzed by using linear mixed models

	Pre ECT (T0; *n* = 65)	During (T1; *n* = 61)	Post ECT (T2; *n* = 65)	Follow up 3M (T3; *n* = 57)	Follow-up 6M (T4; *n* = 51)	Statistic	Corrected *p* value[§]
MoCA Total score	22.46 ± 4.87[†,‡]	22.15 ± 5.22[‡]	21.89 ± 4.89[‡]	23.51 ± 4.70[†]	23.61 ± 4.37[†,‡]	$F(4, 230.5) = 4.14$; $p = .003$	
Subscores:							
- Visuo-executive (/5)	3.74 ± 1.33	3.84 ± 1.42	3.83 ± 1.33	3.95 ± 1.17	4.14 ± 1.08	$F(4, 231.1) = 0.63$; $p = .643$	n.s.
- Naming (/3)	2.89 ± 0.31	2.90 ± 0.35	2.83 ± 0.38	2.91 ± 0.29	2.88 ± 0.32	$F(4, 231.1) = 0.57$; $p = .687$	n.s.
- Attention (/6)	4.88 ± 1.30[†]	4.43 ± 1.85[†,‡]	4.78 ± 1.52[†]	4.45 ± 2.11[†,‡]	4.25 ± 2.34[‡]	$F(4, 251.8) = 3.30$; $p = .012$	$p = .070$
- Language (/3)	1.45 ± 1.02	1.45 ± 0.95	1.49 ± 1.0	1.45 ± 1.09	1.36 ± 1.11	$F(4, 252.1) = 0.44$; $p = .782$	n.s.
Number of Words	**8.86 ± 4.50[†,‡]**	**8.79 ± 4.56[†,‡]**	**8.02 ± 4.10[‡]**	**9.72 ± 4.75[†]**	**10 ± 5.03[†]**	**$F(4, 230.7) = 3.89$; $p = .005$**	**$p = .036$**
- Abstraction (/2)	1.69 ± 0.53	1.75 ± 0.47	1.68 ± 0.50	1.77 ± 0.46	1.73 ± 0.45	$F(4,231.4) = 0.59$; $p = .667$	n.s.
- Delayed recall (/5)	**1.97 ± 1.61[†,‡]**	**1.61 ± 1.54[‡]**	**1.51 ± 1.51[‡]**	**2.21 ± 1.72[†]**	**1.98 ± 1.44[†,‡]**	**$F(4, 231.9) = 3.68$; $p = .006$**	**$p = .045$**
- Orientation (/6)	5.4 ± 1.09	5.36 ± 1.25	5.32 ± 1.19	5.51 ± 1.07	5.75 ± 0.80	$F(4, 231.4) = 1.90$; $p = .112$	n.s.

Data are presented as mean ± standard deviation (SD), reported F-statistic and p-value from multilevel regression models with a compound symmetry covariance structure.

T0 = one week prior to ECT; T1 = during (before the third session); T2 = within 1 week after the index course of ECT; T3 = 3 months after index ECT course; T4 = 6 months after index ECT course; Montreal Cognitive Assessment (MoCA).

[†,‡]: Post hoc correction for multiple testing by using Tukey's HSD: values that do not share the same superscript on the same line are statistically different.

[§]: Significant outcomes after applying Bonferroni–Holm correction for subscores are shown in bold.

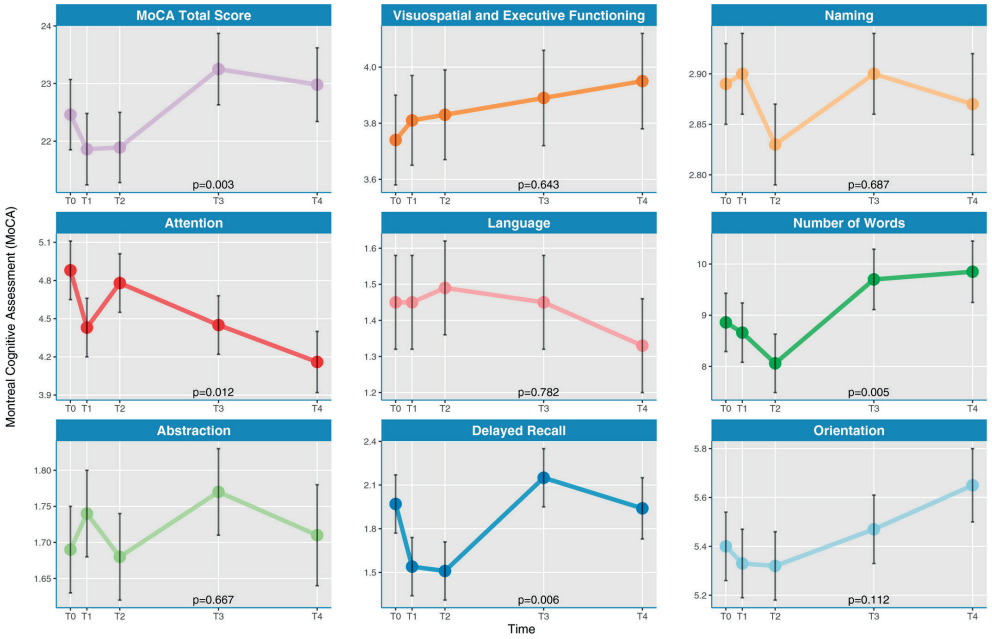


Figure. 2. Change in MoCA total score and subscores. T0 = one week prior to ECT; T1 = during (before the third session); T2 = within 1 week after the index course of ECT; T3 = 3 months after index ECT course; T4 = 6 months after index ECT course; Montreal Cognitive Assessment (MoCA).

Concerning the evolution of MoCA total scores, we found a significant effect of a number of covariates (see Table S1). The three covariates education, age ≥ 65 and number of sessions had a significant main effect, indicating that the MoCA total score at all time points was lower for older patients (main effect - age ≥ 65, $F(1, 63) = 13.27$; $P < 0.001$ estimate: - 3.77 CI [-5.84, -1.70]), for patients with a higher number of ECT sessions (main effect – number of sessions, $F(1, 250.4) = 5.78$; $P = 0.02$; estimate: -0.07 CI [-0.13, -0.01]) and for patients with a lower educational level (main effect – education $F(1, 62.6) = 23.75$; $P < 0.001$; estimate:-4.86 CI [-6.61,-2.76]. The overall evolution of MoCA total scores did not change after adding these three covariates to the model (MoCA at T3> T2, T1). The main effect of HDRS-17 at every time points was significant ($F(1, 248.9) = 10$; $P = 0.002$) and by adding this covariate to the model, we obtained a significant MoCA decrease between T1 and T0 ($P = 0.008$), T2 and T0 ($P = 0.031$) and a significant increase between T3 and T2 ($P = 0.004$). Hence even after adjustment for HDRS we observed a change in cognitive functioning that was independent of the overall improvement in the severity of depression. Two covariates had a significant interaction with time: psychosis (interaction effect $F(4, 226.7) = 2.37$; $P = 0.054$, borderline significance) and baseline cognitive impairment ($F(4, 227.6) = 3.13$; $P = 0.02$) indicating that these patient groups had a different MoCA evolution over time. The non-cognitive impaired group showed a significant decrease from T0 to T2 ($P = 0.041$)

and a significant increase hereafter to T3 (P = 0.013). For the cognitive impaired group, there were no significant differences between the time points. The psychosis group had a significant improvement from T1 to T3 (P = 0.015) and T4 (P = 0.004) and the non-psychosis group had no significant differences between the time points. The evolution of MoCA total score for the confounders age ≥ 65, psychosis and baseline cognitive impairment are represented graphically in Figure S1. The main effect electrode placement was not statistically significant (F(2, 61.6) = 0.86; P = 0.429).

Using the Reliable Change Index to evaluate individual variability in global cognitive functioning between patients

The RCI analyses for the MoCA total score at different time points can be found in Table 3. At T2, 26% of patients (n = 17) had a significantly deteriorated global cognitive functioning compared to T0, 12% (n = 8) improved significantly, and 62% (n = 40) remained stable. At T4, 8% of patients (n = 4) had significantly deteriorated, 18% (n = 9) had significantly improved, and 75% (n = 38) remained stable compared to baseline global cognitive functioning. Demographic and clinical characteristics of patients that deteriorated, improved or remained stable are presented in Table S2: the deteriorators at T2 had a higher MoCA total score at baseline (F(2, 62) = 9.75; P < 0.001).

Table 3: Number of participants who worsened, remained stable, or improved on the MoCA total score compared to baseline according to the Reliable Change Index at the four time points.

	T1	T2	T3	T4
N	61	65	57	51
Worsened *n* (%)	13 (21.3)	17 (26.2)	3 (5.3)	4 (7.8)
Stable *n* (%)	42 (68.9)	40 (61.5)	45 (78.9)	38 (74.5)
Improved *n* (%)	6 (9.8)	8 (12.3)	9 (15.8)	9 (17.6)

Discussion

Using the MoCA, we investigated the impact of ECT on short- and long-term global cognitive functioning in a group of patients with MDE. Our findings show an improvement in global cognitive functioning from T3 compared to T1 and T2, mainly due to an improvement in delayed recall. On an individual level, there was heterogeneity in the effect of ECT on global cognitive functioning; most patients remained stable, but a minority improved or deteriorated significantly.

Our study confirms the findings of previous studies that ECT does not cause persistent global cognitive dysfunction at a group level [5]. However, some important considerations

from previous findings and our own must be addressed. First, because major depression has a known negative influence on cognitive functioning [47], pre-ECT assessment will most likely reflect impaired cognitive abilities. Given this, poor performance at follow-up could be considered to indicate persistent cognitive dysfunction—not compared to pre-ECT (T0) status but compared to predepression functioning. An improvement in mood is expected to induce an improvement in cognitive functioning, but this association could not be confirmed in our study. One possible explanation is that ECT has longer-lasting side-effects that may be masked by improvements coinciding with improvements in mood. Another explanation could be that functional impairment after recovery from depression is a pre-existing vulnerability [48].

Concerning the change in the MoCA subscores, our results are in line with findings from Semkovska andMcLoughlin [5] that show a temporary deterioration in delayed recall and semantic fluency. Contrary to Moirand et al. [14], we found no significant improvement in visuo-executive performance or abstraction MoCA subscores at T2. At the time of this publication, other studies on the long-term change in MoCA scores after ECT do not exist.

We found that patients with psychotic symptoms had lower MoCA scores at all time points, and these patients were particularly vulnerable to cognitive side-effects early in treatment (after 1 week, T1) compared to patients without psychotic symptoms. This last finding contradicts Obbels et al. [49], who found that MMSE scores increased during the whole ECT course and increased proportionally more in psychotic than in non-psychotic depressed patients. These conflicting results could be due to the more accurate measurement by the MoCA of cognitive subdomains such as delayed recall, visuoexecutive functioning and language fluency, which are commonly affected by ECT in patients with MDE [5,24]. However, differences in study sample (elderly versus broad age range) and number of time points (weekly MMSE versus MoCA once during ECT) could also play a role herein.

Although we found no persistent global cognitive dysfunction at the group level, we detected inter-individual variability among adverse cognitive effects of ECT using RCI. This became already apparent 1 week after ECT (T2); 12% of patients improved and 26% significantly worsened in global cognitive functioning. Six months after treatment (T4), the percentage of patients who deteriorated further declined to 8%. The use of RCI analysis in studies of cognitive side-effects of ECT administers relevant supplemental information to group-level results and can provide clinicians in a more nuanced reflection of clinical ECT practice and its associated cognitive side-effects [37,50].

Currently, no 'gold-standard' cognitive screening tool exists to measure ECT-induced cognitive side-effects. Our study confirms the capability of the MoCA to detect both

short- and long-term ECT-induced global cognitive changes in patients with MDE. The MoCA is an easy-to-use test that can be administered by any health professional that has followed the mandatory training. However, the MoCA has a number of drawbacks. First, the MoCA is prone to ceiling effects, especially in younger MDD patients or MDD patients without cognitive impairment [51]. In our study sample, with a mean age of 58.4 and a 41.5 % rate of pre-ECT cognitive impairment, however, the ceiling effects were limited. In younger MDD patients, other global cognitive screening instruments could be considered such as the Screen for Cognitive Impairment (SCIP) [52] which is specifically developed for higher functioning psychiatric patients. Second, the MoCA provides a limited assessment of cognitive subdomains and a more comprehensive test battery is advisable to measure ECT-induced cognitive side-effects. This could consist of the MoCA supplemented with an assessment for retrograde amnesia and a more detailed test for new learning (for example Hopkins Verbal Learning Test-Revised; HVLT-R) [53]. A minimum frequency of three assessments (before, 1 week and 3 months after ECT course) is recommended. Recently, a promising novel cognitive screening tool, the ElectroConvulsive therapy Cognitive Assessment (ECCA) [54] has been proposed which incorporates multiple cognitive assessments (including retrograde amnesia and subjective memory complaints) into one screening tool.

Compared to other studies that have assessed ECT-induced cognitive impairment by using the MoCA, our study is the first to use different versions of the MoCA to investigate long-term MoCA global cognition in a sample of patients with MDE. We also investigated cognitive subdomains and, using the Reliable Change Index, extended the exploration of ECT-induced cognitive impairment by analyzing individual differences. Another strength of this study is the low dropout rate compared to other long-term studies that have assessed cognition after ECT. We maximized efforts to prevent dropout; home assessments were provided when needed.

Our results should also be interpreted in the light of several limitations. Due to the naturalistic design, the sample was heterogenous in terms of demographics (e.g., age) and clinical factors (e.g., psychotic symptoms) that are known to affect cognitive functioning. Approximately half of the patients received C-ECT, which may have confounded results at T3 and T4. Although previous studies suggest that cognitive side-effects are non-progressive under C-ECT [55,56], the administration of CECT may have hampered the recovery of cognitive functioning in the C-ECT group. The evolution of the MoCA total score in the total sample was similar for the subsample that did not receive C-ECT. The RCI analysis of at T3 and T4 should however be interpreted with caution since the small sample sizes of the subgroups preclude us from analyzing the potential confounding role of C-ECT. Second, we did not control for psychotropic drugs during and after ECT treatment. Finally,

the lack of a control group of MDE patients not treated with ECT makes it impossible to conclude whether the described cognitive changes are specific for ECT treatment.

In conclusion, the MoCA is an adequate, but underused, cognitive screener that is able to detect ECT-related cognitive changes in patients with MDE. More research is warranted to investigate the underlying causes of heterogeneity in ECT-induced cognitive impairment.

Supplementary Material

Supplementary Files can be found at: https://onlinelibrary.wiley.com/doi/10.1111/acps.13231

References

1. Geddes J, Carney S, Cowen P et al. Efficacy and safety of electroconvulsive therapy in depressive disorders: a systematic review and meta-analysis. Lancet 2003;361:799-808.
2. Kho KH, van Vreeswijk MF, Simpson S, Zwinderman AH. A meta-analysis of electroconvulsive therapy efficacy in depression. J ECT 2003;19:139–147.
3. Tørring N, Sanghani SN, Petrides G, Kellner CH, Østergaard SD. The mortality rate of electroconvulsive therapy: a systematic review and pooled analysis. Acta Psychiatr Scand 2017;135:388–397.
4. Kumar DR, Han HK, Tiller J, Loo CK, Martin DM. A brief measure for assessing patient perceptions of cognitive side effects after electroconvulsive therapy: the subjective assessment of memory impairment. J ECT 2016;32:256–261.
5. Semkovska M, McLoughlin DM. Objective cognitive performance associated with electroconvulsive therapy for depression: a systematic review and meta-analysis. Biol Psychiatry 2010;68:568–577.
6. Andrade C, Arumugham SS, Thirthalli J. Adverse effects of electroconvulsive therapy. Psychiatr Clin North Am 2016;39:513–530.
7. Ingram A, Saling MM, Schweitzer I. Cognitive side effects of brief pulse electroconvulsive therapy: a review. J ECT 2008;24:3–9.
8. Fink M. Convulsive therapy: a review of the first 55 years. J Affect Disord 2001;63:1–15.
9. Rose D, Fleischmann P, Wykes T et al. Patients' perspectives on electroconvulsive therapy: systematic review. BMJ 2003;326:1363.
10. Waite JEA. The ECT Handbook. Cambridge: Royal College of Psychiatrists, 2013.
11. Porter RJ, Douglas K, Knight RG. Monitoring of cognitive effects during a course of electroconvulsive therapy: recommendations for clinical practice. J ECT 2008;24:25– 34.
12. Thornton A, Leathem J. Cognitive Assessment during a Course of Electroconvulsive Therapy-A National Questionnaire Survey of Current Practice in Aotearoa, New Zealand. Vol 43. 2014.
13. Rasmussen KG. What type of cognitive testing should be part of routine electroconvulsive therapy practice? J ECT 2016;32:7–12.
14. Folstein MF, Folstein SE, White TMM. Mini-mental state examination, 2nd Edition (MMSE-2). Lutz, FL: Psychological Assessment Resources, 2010.
15. Moirand R, Galvao F, Lecompte M, Poulet E, Haesebaert F, Brunelin J. Usefulness of the Montreal Cognitive Assessment (MoCA) to monitor cognitive impairments in depressed patients receiving electroconvulsive therapy. Psychiatry Res 2018;259:476–481.
16. Hoops S, Nazem S, Siderowf AD et al. Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease. Neurology 2009;73:1738–1745.
17. Pinto TCC, Machado L, Bulgacov TM et al. Is the Montreal Cognitive Assessment (MoCA) screening superior to the Mini-Mental State Examination (MMSE) in the detection of mild cognitive impairment (MCI) and Alzheimer's Disease (AD) in the elderly? Int Psychogeriatrics 2019;31:491–504.

18. Fisekovic S, Memic A, Pasalic A. Correlation between moca and mmse for the assessment of cognition in schizophrenia. Acta Inform Med 2012;20:186–189.
19. Blair M, Coleman K, Jesso S et al. Depressive symptoms negatively impact montreal cognitive assessment performance: a memory clinic experience. Can J Neurol Sci 2016;43:513–517.
20. Kalisova L, Kubinova M, Michalec J, Albrecht J, Madlova K, Raboch J. Cognitive functioning in patients treated with electroconvulsive therapy. Neuropsychiatr Dis Treat 2018;14:3025–3031.
21. Seow LSE, Subramaniam M, Chan YWC et al. A retrospective study of cognitive improvement following electroconvulsive therapy in schizophrenia inpatients. J ECT 2019;35:170–177.
22. Tor PC, Ying J, Ho NF et al. Effectiveness of electroconvulsive therapy and associated cognitive change in schizophrenia: a naturalistic, comparative study of treating schizophrenia with electroconvulsive therapy. J ECT 2017;33:272–277.
23. Nasreddine ZS, Phillips NA, B edirian V et al. The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005;53:695–699.
24. Bodnar A, Krzywotulski M, Lewandowska A et al. Electroconvulsive therapy and cognitive functions in treatmentresistant depression. World J Biol Psychiatry. 2016;17:159–164.
25. Rossetti HC, Lacritz LH, Hynan LS et al. Montreal cognitive assessment performance among community-dwelling african americans. Arch Clin Neuropsychol. 2017;32:238–244.
26. Rossetti HC, Lacritz LH, Munro Cullum C, Weiner MF. Normative data for the Montreal Cognitive Assessment (MoCA) in a population-based sample. Neurology 2011;77:1272–1275.
27. Freitas S, Sim~oes MR, Alves L, Santana I. Montreal cognitive assessment: influence of sociodemographic and health variables. Arch Clin Neuropsychol 2012;27:165–175.
28. Roalf DR, Moore TM, Wolk DA et al. Defining and validating a short form Montreal Cognitive Assessment (s- MoCA) for use in neurodegenerative disease. J Neurol Neurosurg Psychiatry 2016;87:1303–1310.
29. Carson N, Leach L, Murphy KJ. A re-examination of Montreal Cognitive Assessment (MoCA) cutoff scores. Int J Geriatr Psychiatry 2018;33:379–388.
30. Krishnan K, Rossetti H, Hynan LS et al. Changes in montreal cognitive assessment scores over time. Assessment 2017;24:772–777.
31. Sofia Costa A, Reich A, Fimm B, Ketteler ST, Schulz JB, Reetz K. Evidence of the sensitivity of the MoCA alternate forms in monitoring cognitive change in early alzheimer's disease. Orig Res Artic Dement Geriatr Cogn Disord 2014;37:95–103.
32. McClintock SM, Choi J, Deng Z-D, Appelbaum LG, Krystal AD, Lisanby SH. Multifactorial determinants of the neurocognitive effects of electroconvulsive therapy. J ECT 2014;30:165–176.
33. Kellner CH, Knapp R, Husain MM et al. Bifrontal, bitemporal and right unilateral electrode placement in ECT: randomised trial. Br J Psychiatry 2010;196:226–234.
34. Verwijk E, Spaans H-P, Comijs HC et al. Relapse and longterm cognitive performance after brief pulse or ultrabrief pulse right unilateral electroconvulsive therapy: a multicenter naturalistic follow up. J Affect Disord 2015;184: 137–144.

35. Squire LR, Chace PM. Squire-1975_Memory 6 months after ECT.pdf. Arch Gen Psychiatry 1975;32:1557–1564.
36. Sackeim HA, Prudic J, Fuller R, Keilp J, Lavori PW, Olfson M. The cognitive effects of electroconvulsive therapy in community settings. Neuropsychopharmacology 2007; 32:244–254.
37. Obbels J, Verwijk E, Vansteelandt K et al. Long-term neurocognitive functioning after electroconvulsive therapy in patients with late-life depression. Acta Psychiatr Scand. 2018;138:223–231.
38. Bjølseth TM, Engedal K, Benth J S, Dybedal GS, Gaarden TL, Tanum L. Clinical efficacy of formula-based bifrontal versus right unilateral electroconvulsive therapy (ECT) in the treatment of major depression among elderly patients: a pragmatic, randomized, assessor-blinded, controlled trial. J Affect Disord. 2015;175:8–17.
39. Sheehan DV, Lecrubier Y, Sheehan KH et al. 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. J Clin Psychiatry. 1998;59:22–33.
40. Trajkovi c G, Star cevi c V, Latas M et al. Reliability of the Hamilton Rating Scale for Depression: a meta-analysis over a period of 49 years. Psychiatry Res 2011;189:1–9.
41. van Diermen L, Hebbrecht K, Schrijvers D, Sabbe BCGG, Fransen E, Birkenh€ager TK. The Maudsley Staging Method as predictor of electroconvulsive therapy effectiveness in depression. Acta Psychiatr Scand 2018;138:605–614.
42. van Diermen L, Versyck P, van den Ameele S et al. Performance of the Psychotic Depression Assessment Scale as a Predictor of ECT Outcome. J ECT 2019;35:238–244.
43. van Diermen L, Vanmarcke S, Walther S et al. Can psychomotor disturbance predict ect outcome in depression? J Psychiatr Res 2019;117:122–128.
44. Thissen AJAM, Van Bergen F, De Jonghe JFM, Kessels RPC, Dautzenberg PLJ. Applicability and validity of the dutch version of the Montreal Cognitive Assessment (moCA-d) in diagnosing MCI. Tijdschr Gerontol Geriatr 2010;41:231–240.
45. Development PG and H-PD. The CORE System. In: Parker G, Hadzi-Pavlovic D, eds. M a, Neurobiological disorder of movement and mood-a phenomenological and review. Cambridge: Cambridge University Press, 1996:82–129.
46. Hinton-Bayre AD. Deriving reliable change statistics from test-retest normative data: comparison of models and mathematical expressions. Arch Clin Neuropsychol. 2010;25:244–256.
47. Rock PL, Roiser JP, Riedel WJ, Blackwell AD. Cognitive impairment in depression: a systematic review and metaanalysis. Psychol Med 2014;44(10):2029–2040.
48. Bos EH, Ten Have M, van Dorsselaer S, Jeronimus BF, de Graaf R, de Jonge P. Functioning before and after a major depressive episode: pre-existing vulnerability or scar? A prospective three-wave population-based study. Psychol Med 2018;48:2264–2272.
49. Obbels J, Vansteelandt K, Verwijk E et al. MMSE changes during and after ECT in late-life depression: a prospective study. Am J Geriatr Psychiatry 2019;27:934–944.

50. Dybedal GS, Tanum L, Sundet K, Gaarden TL, Bjølseth TM. Cognitive side-effects of electroconvulsive therapy in elderly depressed patients. Clin Neuropsychol 2014;28: 1071–1090.
51. Ragguett RM, Cha DS, Kakar R, Rosenblat JD, Lee Y, McIntyre RS. Assessing and measuring cognitive function in major depressive disorder. Evid Based Ment Health 2016;19:106–109.
52. Purdon S. The Screen for Cognitive Impairment in Psychiatry (SCIP): Administration Manual and Normative Data. Edmonton, AB: PNL Inc; 2005.
53. Benedict RHB, Schretlen D, Groninger L, Brandt J. Hopkins verbal learning test – revised: normative data and analysis of inter-form and test-retest reliability. Clin Neuropsychol 1998;12(1):43–55.
54. Hermida AP, Goldstein FC, Loring DW et al. ElectroConvulsive therapy Cognitive Assessment (ECCA) tool: a new instrument to monitor cognitive function in patients undergoing ECT. J Affect Disord 2020;269:36–42.
55. Petrides G, Tobias KG, Kellner CH, Rudorfer MV. Continuation and maintenance electroconvulsive therapy for mood disorders: review of the literature. Neuropsychobiology 2011;64:129–140.
56. Kirov GG, Owen L, Ballard H et al. Evaluation of cumulative cognitive deficits from electroconvulsive therapy. Br J Psychiatry. 2016;208:266–270.

CHAPTER 8

Cognitive trajectories after electroconvulsive therapy in mood disorders

Hebbrecht K, Dejaeger M, Giltay EJ, Birkenhäger TB, Sabbe B, Verwijk E, Obbels J, Schrijvers D, Van Diermen L. .

Journal of Psychiatric Research. 2022 Dec; 156; 132-140

Abstract

Cognitive function during an ECT care pathway is mainly investigated at the group level by analyzing mean cognitive test scores over time. However, there are important inter-individual differences, with some patients experiencing residual invalidating cognitive deficits. This study provides a nuanced examination of cognitive functioning during and after ECT by combining three approaches for data analysis. A cognitive test battery was assessed in seventy-three ECT-treated patients with a Major Depressive Episode (MDE) at up to five time points (baseline, immediately prior to the third session and 1 week, 3 months and 6 months after completion of the index course). Group-level changes in cognitive function were investigated using linear mixed models and individual-level changes were examined using Reliable Change Indices (RCI). The presence of patient subgroups with similar cognitive trajectories was explored using Latent Class Growth Analysis (LCGA). At the group level, there was a temporary deterioration in processing speed, verbal memory and retrograde amnesia during and after index course of ECT. Individual-level analyses revealed considerable variability in cognitive effects of ECT. Three patient classes with a similar cognitive trajectory could be identified, all with a rather parallel courses over time, thus mainly differing in terms of pre-ECT cognitive functioning.

Introduction

Electroconvulsive therapy (ECT) is the most effective biological treatment for patients with severe Major Depressive Disorder (MDD) [1,2]. Although previous research did not show persisting cognitive side-effects after ECT at the group-level [3,4], there is a high level of inter-individual variability in cognitive function after ECT [5,6]. ECT-induced cognitive side-effects can be subdivided into following categories: acute postictal disorientation, anterograde amnesia for recently learned information and retrograde amnesia for previously learned information [7] . Non-memory cognitive functions that can be affected include processing speed, attention, verbal fluency and executive function [3,8,9].

Previous studies mainly investigated cognitive function during an ECT pathway by analyzing changes in mean test scores of the total study sample. Change in cognitive function after ECT is hereby considered as a homogeneous phenomenon across patients [10,11]. Two previous studies have investigated the inter-individual differences in cognitive effects of ECT using Reliable Change Indices. Both studies assessed cognition in depressed elderly at 2 time intervals after ECT and they confirmed a significant cognitive deterioration or improvement in a minority of patients, both at short-term as well as long-term [5,6]. Group-level analyses not only overlook the inter-individual variability in cognitive function, they also overlook possible differences in cognitive trajectories over time among individual patients [12]. We are not aware of previous studies that have investigated the presence of distinct subgroups of ECT treated patients with a similar cognitive trajectory over time during the ECT care pathway and the differentiating characteristics of these groups. Though, this approach would enable to gain more insight into the longitudinal course of cognitive functioning during an ECT care pathway. Furthermore, from a clinical perspective, it would pave the way for identifying high-risk patients and for customizing cognitive assessments according to the patient's risk profile. Moreover, patients and their relatives could be better informed on the time course of cognitive function after ECT. Latent Class Growth Analysis (LCGA) is a promising method for the identification of homogeneous meaningful subgroups, according to their longitudinal pattern, within a larger heterogeneous patient population and is increasingly being used in psychiatry [13,14].

Methods

Study sample

Patients between 18 and 85 years old and diagnosed with a major depressive episode (MDE) (i.e., Major Depressive Disorder or bipolar depression according to the DSMIV-TR) were included. They were recruited from the in- and outpatient departments of the University Psychiatric Hospital in Duffel (Belgium). Diagnoses were confirmed by using

the MINI diagnostic interview version 6.0.3 [15] and a minimum score of 17 out of 52 (i.e., moderate severity) (Zimmerman et al., 2013) on the Hamilton Depression Rating Scale-17 (HRSD-17) [16,17]. Exclusion criteria were a history of substance abuse (< 6 months prior) or a diagnosis of a primary psychotic or schizoaffective disorder. We did not systematically assess the presence of a comorbid neurocognitive disorder (or intellectual disability) as its symptoms can be hard to differentiate from those of an MDE. The data are part of the PROTECT study [18–20], which was designed to investigate ECT response predictors. All patients provided written informed consent before the study procedures were performed. The study protocol complied with the Declaration of Helsinki and was approved by the local medical ethics committee. The study is registered in the online clinical database ClinicalTrials.gov (Identifier: NCT02562846).

Cognitive assessments

The cognitive test battery consisted of four validated and well-established cognitive tasks. The Montreal Cognitive Assessment (MoCA) [21] was used to assess global cognitive function, the results of which are discussed more thoroughly in a previous article [22]. The MoCA consists of thirteen subtasks that assess eight cognitive domains (visuospatial, visuo-executive, naming, attention, language, abstraction, delayed recall, and orientation). The time to administer is approximately 10 minutes and the maximum score is 30, obtained by summing the subscores [21]. The Symbol Digit Substitution Test (SDST) [23] was performed to test processing speed. In this task, a series of symbols is presented that need to be matched with corresponding numbers as quickly as possible according to a key provided at the top of the page. The Hopkins Verbal Learning Test-Revised (HVLT-R) [24] was done to test verbal episodic memory. The test consists of three learning trials in which 12 nouns within 3 semantic groups are read aloud by the administrator. The outcome measure HVLT-R Total Learning (HVLT-R TL) refers to the sum of reproduced words at every trial and the HVLT-R Delayed Recall (HVLT-R DR) to the number of words that were reproduced 20 minutes after the last learning trial. Section C (AMI-C) of the Kopelman Autobiographic Memory Interview [25] was used to measure retrograde amnesia. Section C, which investigates memories over the past year (the recent period), was chosen because the recent memories are most frequently impaired after ECT [26]. The MoCA, HVLT-R and SDST were assessed at five time points: in the week prior to ECT (T0), during ECT (before the third session, T1), within 1 week after completion of the index course (T2) and at 3 (T3) and 6 (T4) months after T2. The AMI-C was assessed at the same intervals except for T1. When patients were not able to come to the clinic at T3 and T4, they were tested at their homes. To minimize learning effects, we used different versions of the MoCA, HVLT-R and SDST. Participants that were not able to complete the full cognitive test battery, completed a limited protocol (i.e., minimum assessment of the MoCA, if possible supplemented with one or more of the other cognitive tasks).

Mood assessments

The severity of depressed mood was assessed with the Hamilton Rating Scale for Depression-17 (HRSD-17) [16,17] at each time point (T0–T4) as well as weekly during the ECT course. The HRSD-17 is one of the most used instruments for assessing mood severity and response to treatment, both in clinical practice as well as for research purposes. We additionally used the Montgomery-Asberg Depression Rating Scale (MADRS) [27], at the same time intervals, as this scale is more sensitive to treatment effects compared to the HRSD-17.

ECT procedure

Patients received a course of twice weekly ECT using a brief-pulse (0.5 milliseconds), constant-current Thymatron IV system (Somatics LLC, Lake Bluff, Illinois). Electrodes were placed unilaterally over the right hemisphere or bilateral when a fast symptom reduction was required. Based on international guidelines [28], patients were switched from unilateral to bilateral ECT in case of lack of treatment response after six treatments. We defined the lack of response as a decrease of less than 25% of the HRDS-17 score, in addition to the appraisal of clinical response by an experienced psychiatrist. The age-based dosing method was used to determine the ECT stimulus for unilateral treatment and the half age-based method in case of a bilateral electrode placement. Anesthesia was achieved with etomidate (0.15 mg/kg) or propofol (1 mg/kg) in case of intolerance to etomidate. When the overall clinical effect of ECT was not significant after 12 sessions, ketamine (1-2 mg/kg) was used as an induction agent) [29]. Succinylcholine (0.5 mg/kg) was provided for muscle relaxation. The ECT sessions were terminated in case of remission or lack of further clinical improvement in the last three sessions.

Statistical analysis

Baseline demographic and clinical characteristics are presented using means and standard deviations for, normally distributed, continuous variables and frequencies for categorical variables.

First, we analyzed cognitive functioning of the whole group over time using LMM. The raw scores on the different cognitive tasks were used as outcome, subject was included as random intercept effects and time (time points T0 through T4) as a fixed effect. Since we had six different cognitive task scores as outcome variables (MoCA, HVLT-R TL, HVLT-R DR, SDST, AMI-C autobiographical and AMI-C personal semantic), we applied a Bonferroni correction for multiple testing.

In a second step, we investigated the variability in cognitive effects at individual patient-level by examining the reliable change index (RCI) at different time points [30]. Specifically, when group-level analyses show only a small (nonsignificant) average change in cognitive

function, RCI enables to detect the subgroup of patients with significant cognitive decline or improvement after ECT. The RCI was calculated using the method proposed by Jacobson and Truax—RCI = (posttest - pretest)/Standard Error (SE) [31]. A 90% confidence interval was used to indicate a significant change. Thus, RCI values of 1.645 or higher were considered as a "reliable improvement" and values of – 1.645 or lower as "reliable decline" [32]. The SE was calculated using the variance in baseline scores of a control sample and the reliability of the test extracted from psychometric papers of the respective test [21,24,33,34]. No normative scores for the part C of the AMI were found so the AMI-C was excluded from the RCI analyses.

Lastly, we used LCGA to identify subgroups of patients with similar cognitive trajectories. Six composite (sub)scale scores on cognition (MoCA, HVLT-R TL, HVLT-R DR, SDST, AMI-C autobiographical and AMI -C personal semantic), were all group-standardized, after which these were averaged to yield the pooled standardized cognitive dysfunction, which was used as the outcome in the LCGA analysis. LCGA classifies individual patients into classes based on their similarity of model parameters across varying time points. As it is impossible to definitively determine a patient's class membership, there is a probability attached to each patient with respect to the appropriate class. The fit measures included the Bayesian information criterion [BIC], Akaike information criterion [AIC] and size-adjusted BIC [saBIC]. Entropy was used to assess the classification accuracy with high values (>.80) indicating a good discriminatory power [11]. The lower the AIC, BIC, and saBIC, the more superior the fit (shown in elbow plots in the Supplementary Figure 1). To determine what improvement is made to the model by adding an additional class, the differences between the values for the AIC, BIC, and saBIC were calculated. If the difference between the values of one additional class was small, an additional class was considered to have limited additional value. Models were estimated using Maximum Likelihood estimation, using each participant's non-missing data. One-hundred random starts were used to prevent solutions at local maxima, resulting in the same classes. As age has a major determining influence on cognitive function, we added age as a confounder to the model. This holds the advantage of finding the latent classes that were independent of the impact on age on the longitudinal trajectories of cognition of patients.

Hereafter, the three trajectories were compared with each other according to nine potential predictor variables (i.e., sex, educational level, mood disorder type, HRSD and MoCA score, presence of psychotic and melancholic symptoms, bilateral electrode placement and placement switch) using univariate multinomial logistic regression analyses, corrected for age.

Additionally, we plotted the course of the HRSD-17 and the MADRS over time for the three trajectories to investigate the role of the evolving variable of depression on cognitive trajectories.

All statistical analyses were performed using R version 3.4.3 ("R: A language and environment for statistical computing. R Foundation for Statistical Computing; 2019. https://www.R-project.org/. Accessed 31 August 2020.," n.d.) using RStudio (R Foundation for Statistical Computing, Vienna, Austria, 2016. URL: https://www.Rproject.org/), with main packages 'nnet' for the multinomial log-linear models (version 7.3-12) and 'lcmm' for the LCGA (version 1.8.1). Significance was based on a twotailed < 0.05, except for the univariate multinomial logistic regression analysis, we applied a significance rate of < 0.01 (99% CI).

Results

Demographic and clinical characteristics

The demographic, clinical and treatment characteristics of the participants are shown in Table 1. Seventy-three patients were included in the study. Premature termination of ECT occurred in eight patients and this was due to intolerable side effects (n = 5, of which n = 1 due to cognitive effects) or other, non-treatment related, reasons (n = 3). Patients had a mean age of 59 years old, and nearly half of the patients had psychotic symptoms at baseline. Thirty-two patients (43.8%), mainly aged over 65 years, had a global cognitive impairment at baseline (i.e., MoCA <23) (Carson et al., 2018). Fiftyfour out of 73 patients (73.9 %) reached response and 41 (56.2%) reached remission after ECT. Thirty-five (47.9%) patients received continuation ECT (C-ECT) during the 6-month time interval after completing the index course.

Table 1. Demographic, clinical and treatment characteristics of the study population

N	**73**
Age, years mean (± SD); range	58.8 (± 15.1); 21–85
Gender, female *n* (%)	56 (76.7)
Level of education > 12 n (%)	38 (52.1)
Bipolar *n* (%)	13 (17.8)
Psychotic features *n* (%) at baseline	33 (45.2)
Melancholic features (CORE-defined) *n* (%)	46 (63.0)
Episode duration (Months) - Mean (± SD) - Median, range	 14.3 (± 18.1) 6.5, 1–84
HDRS-17 mean (± SD) baseline (T0); range	24.8 (± 6.0); 17–38
Responders at T2 *n* (%)	54 (73.9)
Remitters at T2 *n* (%)	41 (56.2)
Cognitive impairment at baseline (MoCA < 23) *n* (%)	32 (43.8)
Number of ECT treatments in acute course, mean (± SD); range	11.2 (± 5.7); 2–27
Patients who received continuation ECT *n* (%)	35 (47.9)
Electrode placement - Right unilateral *n* (%) - Bilateral *n* (%) ○ Bifrontal *n* (%) ○ Bitemporal *n* (%) - Mixed (Switch) *n* (%)	 41 (56.2) 13 (17.8) 2 (2.7) 11 (15.1) 19 (26.0)

Note. CORE Assessment of psychomotor functioning (CORE), Hamilton Depression Rating Scale (HDRS), Montreal Cognitive Assessment (MoCA), Electroconvulsive Treatment (ECT)

Evolution of cognitive functioning on group-level

The evolution of the different cognitive assessments over time, as analyzed using LMM, are shown in Figure 1 and Supplementary Table 1. As can be seen from Supplementary Table 1, not all patients were able to complete the full cognitive test battery and these patients completed a limited protocol (i.e., MoCA and if possible supplemented with one or more of the other cognitive tasks). As shown in Supplementary Table 1, 77% (56/73) patients completed the full cognitive test battery at T0 and this percentage dropped to 74% (54/73) at T2 and 56 (41/73) at T4. The MoCA showed a significant improvement at T3 compared to T1 and T2 and at T4 compared to T2 ($F(4, 253.3) = 5.32$; $p = .0016$). The HVLT-R-TL score was significantly higher at T3 and T4 compared to T1 and T2, there were no significant differences in pairwise comparisons with T0 ($F(4,197.0) = 6.12$; $p = .0006$). The HVLT-R-DR significantly decreased at T2 compared to T0 but improved towards baseline levels at T4 ($F(4,195.9) = 6.70$; $p = .0006$). The number of correct digits reproduced on the SDST significantly decreased at T1 compared to T0 but improved to levels comparable to T0 at time points T3 and T4 ($F(4, 195.9) = 6.29$; $p < .0001$). The AMI-C personal semantic

score was lower at T2 compared to T0 (F(3, 142.0) = 5.54; p = .0039). At T4, the score had improved to baseline levels.

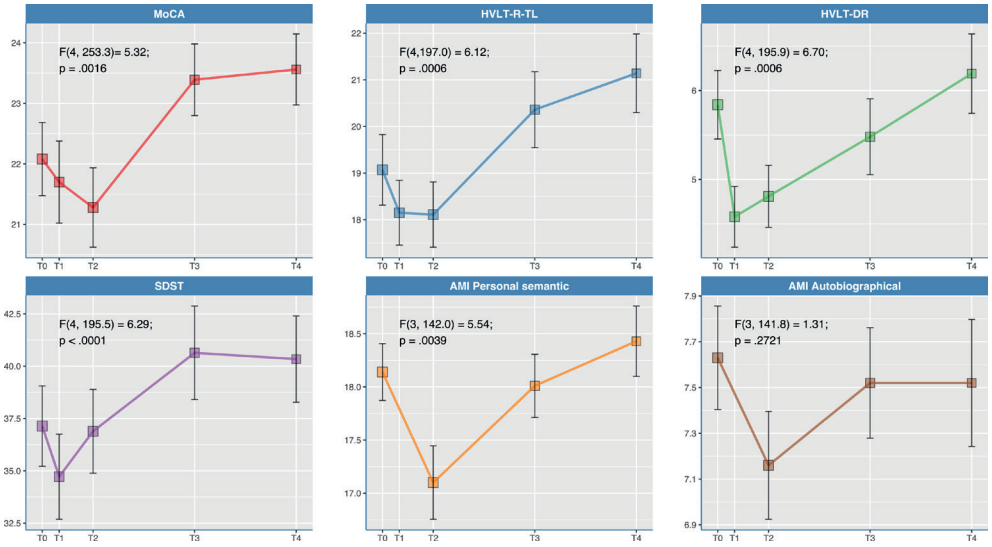


Figure 1: Change in cognitive test scores over time

Note. Data are presented as mean ± standard deviation (SD) as analyzed by using linear mixed models, reported F-statistic and p-value from multilevel regression models. P-values are Bonferroni-Holm corrected.

Abbreviations. T0 = one week prior to ECT; T1 = during (before the third session); T2 = within 1 week after the index course of ECT; T3 = 3 months after index ECT course; T4 = 6 months after index ECT course; Montreal Cognitive Assessment (MoCA); Hopkins Verbal Learning Test-Revised Total Learning (HVLT-R TL); Hopkins Verbal Learning Test-Revised Delayed Recall (HVLT-R DR); Symbol Digit Substitution Test (SDST); Autobiographic Memory Interview Section (AMI-C).

Individual differences in cognitive functioning

Table 2 shows the RCI analyses for the different cognitive tests at the four time points. AMI-C was not included in the analyses because there were no normative data available. Although two cognitive tests (MoCA, HVLT-R-TL) did not show significant group-level differences compared to baseline (T0), RCI analyses revealed some individual differences with a subgroup of patients (20 to 26 %) deteriorating, improving (7 to 11 %) or showing no differences post-ECT (T2) compared to baseline (T0) (63-73 %). The mean (group-level) scores of the HVLT-R-DR and SDST showed a significant decrease post-ECT compared to baseline, but RCI analyses revealed subgroups of patients that improved or showed no differences compared to baseline (for an overview of all percentages, see Table 2). At T4, the percentages of improvers for all cognitive test scores were consistently higher than the percentages of deteriorators.

Table 2: Effect of ECT on individual levels of cognitive functioning †

	During (T1)	Post ECT (T2)	Follow up 3M (T3)	Follow-up 6M (T4)
MoCA				
- Worsened	14 (20.3)	19 (26.4)	4 (6.5)	4 (7.4)
- Stable	49 (71.0)	45 (62.5)	49 (79.0)	41 (75.9)
- Improved	6 (8.7)	8 (11.1)	9 (14.5)	9 (16.7)
HVLT-R-TL				
- Worsened	10 (19.2)	11 (20.0)	3 (6.0)	1 (2.3)
- Stable	41 (78.9)	40 (72.7)	41 (82.0)	36 (83.7)
- Improved	1 (1.9)	4 (7.3)	6 (12.0)	6 (14.0)
HVLT-R-DR				
- Worsened	16 (30.8)	13 (24.5)	9 (18.0)	4 (9.3)
- Stable	31 (59.6)	37 (69.8)	36 (72.0)	33 (76.7)
- Improved	5 (9.6)	3 (5.7)	5 (10.0)	6 (14.0)
SDST				
- Worsened	12 (22.6)	7 (12.7)	4 (8.0)	5 (12.2)
- Stable	40 (75.5)	42 (76.4)	39 (78.0)	30 (73.2)
- Improved	1 (1.9)	6 (10.9)	7 (14.0)	6 (14.6)

Data are n (%).
† No normative data for part C of the AMI were found.

Differential cognitive trajectories after ECT

In order to examine the presence of distinct trajectories of cognitive function after ECT, we performed LCGA analyses by grouping patients according to their longitudinal trajectory (using a six composite cognition score as outcome; i.e. including the MoCA, HVLT-R TL, HVLT-R DR, SDST, AMI-C autobiographical and AMI-C personal semantic). Table 3 shows the estimation process and the corresponding fit indices for each model. The three-trajectory model was selected as the best fit. The AIC and (SA)-BIC showed a significant drop from a three to four-trajectory model and hereafter the differences were significantly smaller. Additionally, in the four-trajectory model, some trajectories included less than 5 patients (6.9%) (Table 3 and Supplementary Figure 1). The three cognitive trajectories are shown in Figure 2A.

Table 3. Fit Indices of One- to Six Class Latent Growth Mixture Models over a Six-Month Follow-up

N classes	Maximum Likelihood	AIC	BIC	SA-BIC	Entropy	Percentage of Individuals in Class					
						1	2	3	4	5	6
1	-392.5	793.0	802.2	789.5	1.00	100					
2	-318.8	651.5	667.5	645.5	0.96	17.81	82.19				
3	-281.8	583.6	606.5	575.0	0.89	10.96	45.21	43.84			
4	-271.2	568.4	598.2	557.2	0.88	6.85	10.96	41.10	41.10		
5	-268.43	568.9	605.5	555.1	0.80	6.85	10.96	35.62	31.51	15.07	
6	-268.43	574.9	618.4	558.5	0.77	6.85	0	10.96	32.88	35.62	13.70

Note. AIC = Akaike Information Criterion; BIC Bayesian Information Criterion; SA-BIC = size-adjusted BIC

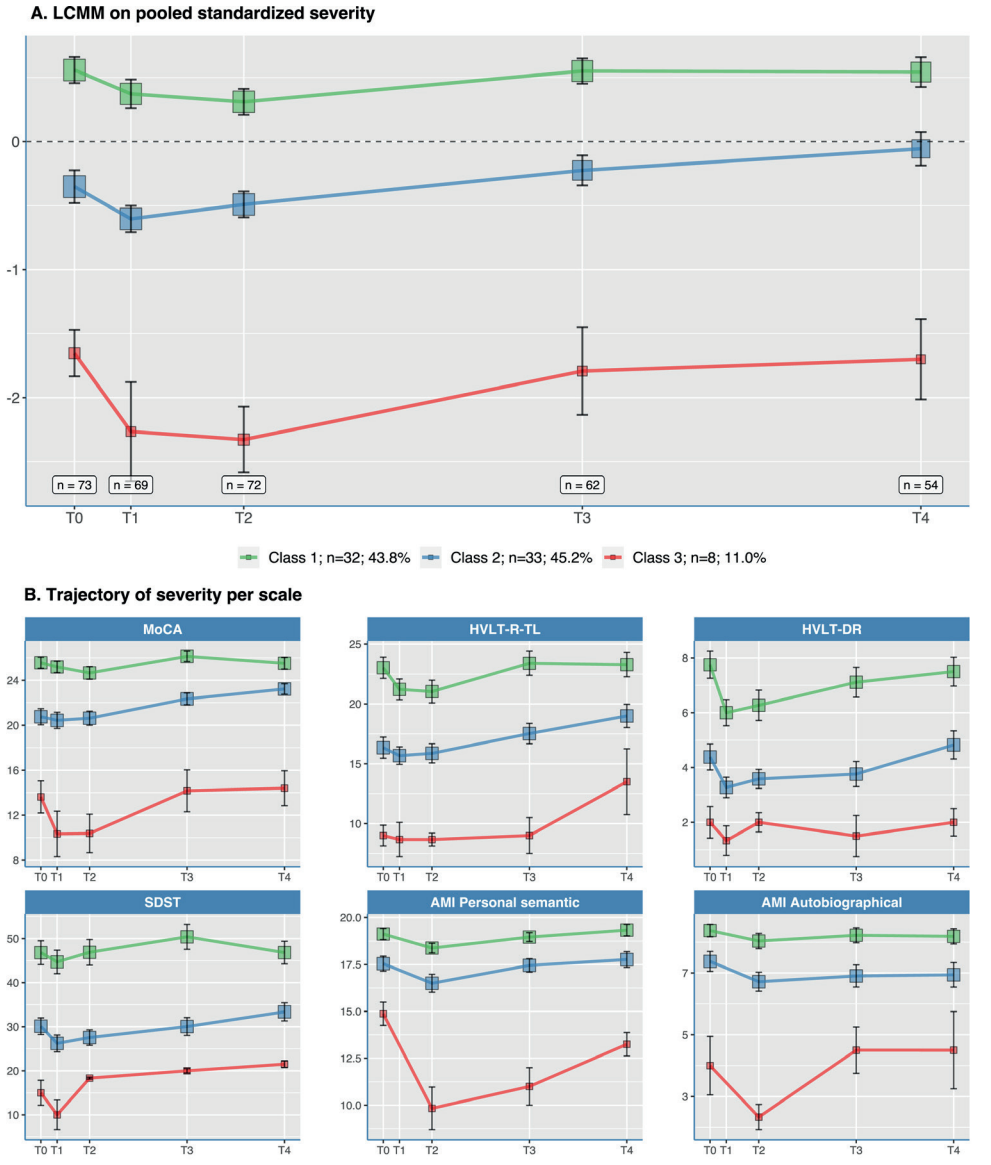


Figure 2. Longitudinal Cognition Trajectories

Note. Longitudinal LCGA-based trajectories are derived from on a composite cognitive (sub)task score of six tasks (MoCA, HVLT-R TL, HVLT-R DR, SDST, AMI autobiographical and AMI personal semantic).

Abbreviations. T0 = one week prior to ECT; T1 = during (before the third session); T2 = within 1 week after the index course of ECT; T3 = 3 months after index ECT course; T4 = 6 months after index ECT course; Montreal Cognitive Assessment (MoCA); Hopkins Verbal Learning Test-Revised Total Learning (HVLT-R TL); Hopkins Verbal Learning Test-Revised Delayed Recall (HVLT-R DR); Symbol Digit Substitution Test (SDST); Autobiographic Memory Interview Section (AMI-C).

The three trajectories primarily differ from each other by their baseline (pre ECT) cognitive functioning. The first (Class 1) trajectory, including 32 patients (43.8%), showed a good cognitive performance at baseline, followed by a decline during and after ECT and a recovery towards pre ECT cognitive functioning. The second (Class 2) trajectory, including 33 patients (45.2%), showed a moderately impaired cognitive performance at baseline and a comparable course to Class 1 (mild decline and improvement thereafter). The third (Class 3) trajectory, including only 8 patients (11.0%), had a baseline impaired cognitive function and, likewise, a comparable course to Class 1 and 2.

Fig 2B shows the cognitive trajectories per cognitive task of patients in Class 1 to 3. The general tendency of the task trajectories was similar to that of the respective overall cognitive trajectories (of Class 1 to 3).

Clinical predictors for the cognitive trajectories

Univariate analyses, corrected for age, showed that a low level of educational, a low baseline MoCA score and the presence of melancholic features at baseline were significant predictors for membership to Class 2 or 3 versus Class I as the reference category (Figure 3: Low education: Class 2 vs. 1: OR = 2.51, $p < 0.001$, 3 vs. 1: OR = 5.05, $p < 0.005$; baseline MoCA: Class 2 vs. 1: OR = 9.62, $p < .001$; 3 vs. 1: OR = 72.0, $p < .0001$; melancholic features (Class 2 vs. 1: OR = 1.97, $p = .01$; 3 vs. 1: OR = 2.91; $p = .051$). As participants belonging to Class 3 were exclusively female, the confidence interval for Class 3 vs. 1 was not estimable. The other investigated factors (baseline HRSD, mood disorder type, Psychotic features, bilateral placement, placement switch) were no significant predictors for membership to Class 2 or 3.

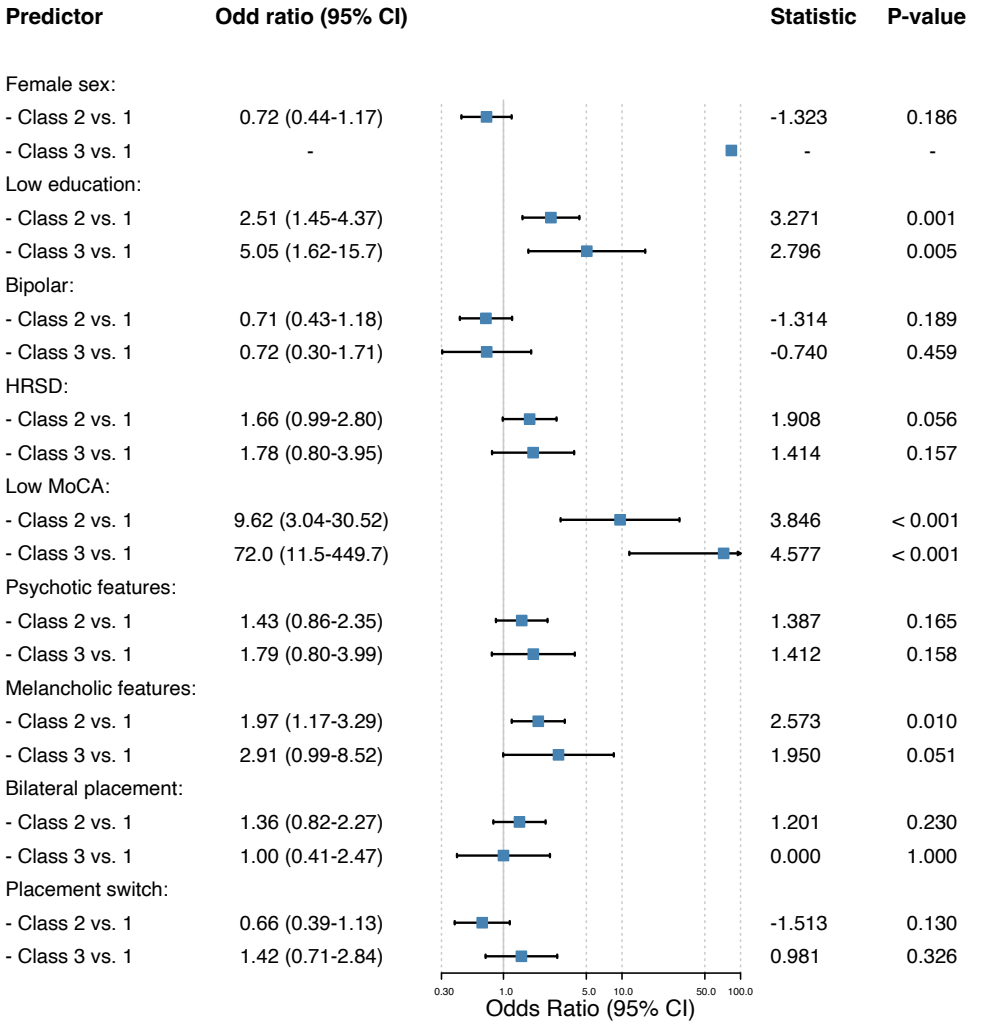


Figure 3. Forest plot for univariate multinomial logistic regression
Abbreviations. Hamilton Rating Scale for Depression (HRSD); Montreal Cognitive Assessment (MoCA).

Evolution of cognitive task scores along with depressive symptomatology

Figure 4 depicts the evolution of severity of depressive symptoms during ECT for the three cognitive trajectory groups over time. All three groups show a significant improvement in mood over time which stabilized directly after ECT for Class 1 and 2. Class 3 showed further mood improvement thereafter. As cognitive functioning during and after ECT showed a decline but mood improved during that same time interval, the (temporary) decline in cognitive function could not be fully explained by the differences in improvement of depressive symptom severity.

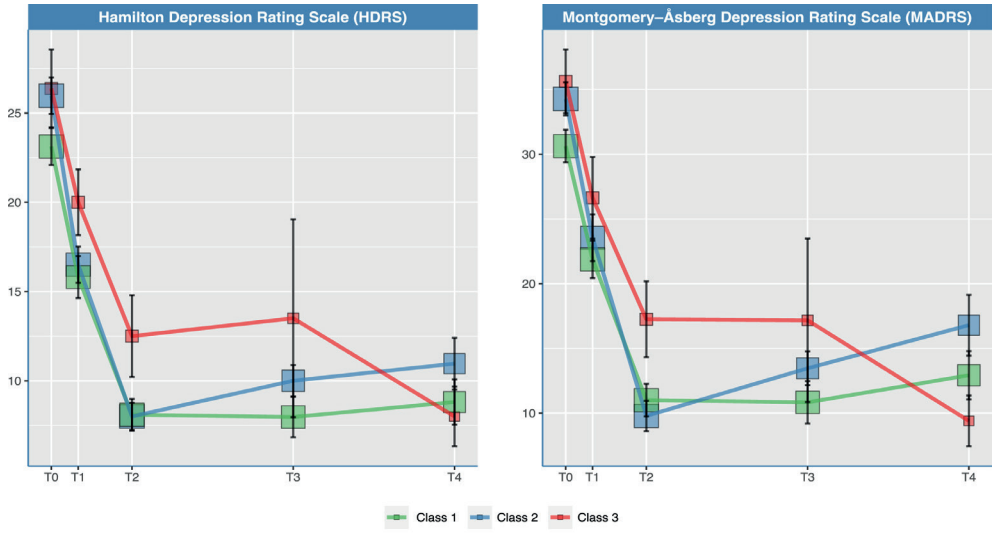


Figure 4: Evolution of mood according to the three cognitive trajectories

Discussion

In this study, we provided a nuanced investigation of cognitive function during and after ECT by combining three approaches. First, cognitive function was investigated by studying mean group-level outcomes on cognitive test scores. Second, individual-level outcomes were examined and, third, the presence of subgroups of patients with a similar cognitive trajectory was explored.

Our group-level analyses showed a temporary deterioration during and directly after ECT completion in average test scores of processing speed, verbal episodic memory (learning and delayed recall) and retrograde (personal semantic) memory. There were no significant changes in global cognitive impairment (as measured by the MoCA) compared to baseline. All cognitive task scores returned to baseline values at long-term follow-up. These reassuring findings are consistent with the results from the meta-analysis by Semkovska e.a. [3] indicating that average cognitive scores all return to baseline levels or normalize at long-term follow-up (weeks to months after acute ECT session) [3,4,35] Our results did not support the findings of group-level persistent impairment of autobiographical memory after ECT in previous studies [36–38]. A detailed study of the MoCA evolution, including the evolution of the subscores is provided in our previous manuscript [22].

The individual-level analyses supplemented the group-level findings by providing a more fine-grained perspective on cognitive function. The individual-level analyses confirm

that cognitive function after ECT is a highly individual phenomenon with subgroups of patients showing either cognitive deterioration, improvement or no change compared to baseline at short- and/or long-term follow-up. These findings are in accordance with previous studies in elderly patients with depression receiving ECT [5,6]. This inter-individual variation is possibly to be explained by a variation and interaction between individual vulnerability and ECT technical parameters [39]. It is possible that the subgroup of patients experiencing cognitive side effects include patients with an undiagnosed comorbid neurocognitive disorder (or a prodrome thereof).

Our study not only examined the variability in cognitive function on different time points, the variability in longitudinal trajectories of cognitive function over time was also explored. Patients could be classified into three subgroups according to their cognitive trajectory. None of the subgroups was characterized by an overall marked deterioration or improvement in their cognitive functioning. The three trajectories rather showed a fairly parallel course characterized by an initial decline in cognitive function followed by a stabilization towards baseline levels thereafter. So, the cognitive function after ECT is a highly individual phenomenon but the trajectory itself do not appear to be highly different between individuals. Our results however should be interpreted with caution considering the small sample size of the subgroups, and therefore these findings need to be confirmed in other larger groups of patients undergoing ECT. Importantly, we found that trajectories of cognitive function were mainly determined by baseline characteristics (i.e., educational level, baseline MoCA and presence of melancholic features) but not by ECT treatment or mood improvement. Age could not be included as a predictor for membership to a specific trajectory as we adjusted the LCGA analyses for age. Nevertheless, age showed a consistent inverse correlation with cognitive scores at every time point (see Supplementary Figure 2), suggesting that cognitive recovery after ECT was not determined by age.

Why cognitive dysfunction persists in a subgroup of patients successfully treated with ECT is not clear. The absence of premorbid cognitive scores hampers our interpretation. As already stated by Obbels e.a. [6], one explanation could be that depression leaves a 'cognitive scar' in a subgroup of patients. Another explanation could be that cognitive dysfunction itself is a trait marker that increases the risk for depression in some patients. RoCa e.a. [40] similarly found persistence of cognitive dysfunction in remitted patients with melancholic depression, after pharmacological treatment, at 6 months follow-up and suggested that this cognition dysfunction could represent a trait marker of the melancholic depression subtype.

Our study has some important strengths. First, the repeated measurements of several cognitive domains over a period of 6 months allow a detailed insight into the longitudinal evolution of cognition over time. Second, different versions of the cognitive tests were

used at subsequent measurement points and all assessments were done by the same psychiatrist-researcher (LVD), excluding interrater bias. Third, in comparison with other cognition ECT studies, there was a limited drop-out of patients.

There are also some noteworthy limitations to our study. First, although previous studies imply that cognitive side effects are non-progressive under C-ECT [6,41,42], the administration of C-ECT may have influenced results on cognitive test scores at T3 and T4. Second, not all patients were able to complete all cognitive tests at all time points and this might introduce potential selection bias. Third, although an extensive cognitive test battery was used, it lacks a thorough assessment of executive function which is also shown to be affected after ECT [3]. Furthermore, the percentage consistency of autobiographical memory is not assessed by the AMI-C, which makes it less suitable for repeat assessments. Fourth, the relatively small number of patients precluded multivariate logistic regression analyses in the prediction model of the trajectory classes. Further research with larger sample sizes is required for validation of the cognitive trajectories and their predictive variables. Furthermore, as the effect of ECT on cognitive performance is a complex clinical phenomenon with multiple determinants, future studies regarding pretreatment moderators of cognitive trajectories, should investigate a wide range of demographic, clinical and neurobiological variables, in addition to the nine variables included in our study.

Our study points towards the inter-individual variability in cognitive function during an ECT care pathway and highlights the importance of cognitive monitoring during and after treatment. As a small group of patients shows persisting cognitive impairment, with a potentially great impact on psychosocial functioning, early recognition of this minority group is highly important. Our study adds to the evidence for a more person-oriented cognitive screening approach and states that patients with lower baseline MoCA scores or presence of melancholic features merit increased cognitive monitoring (more extensive battery and higher frequency).

Supplementary Material

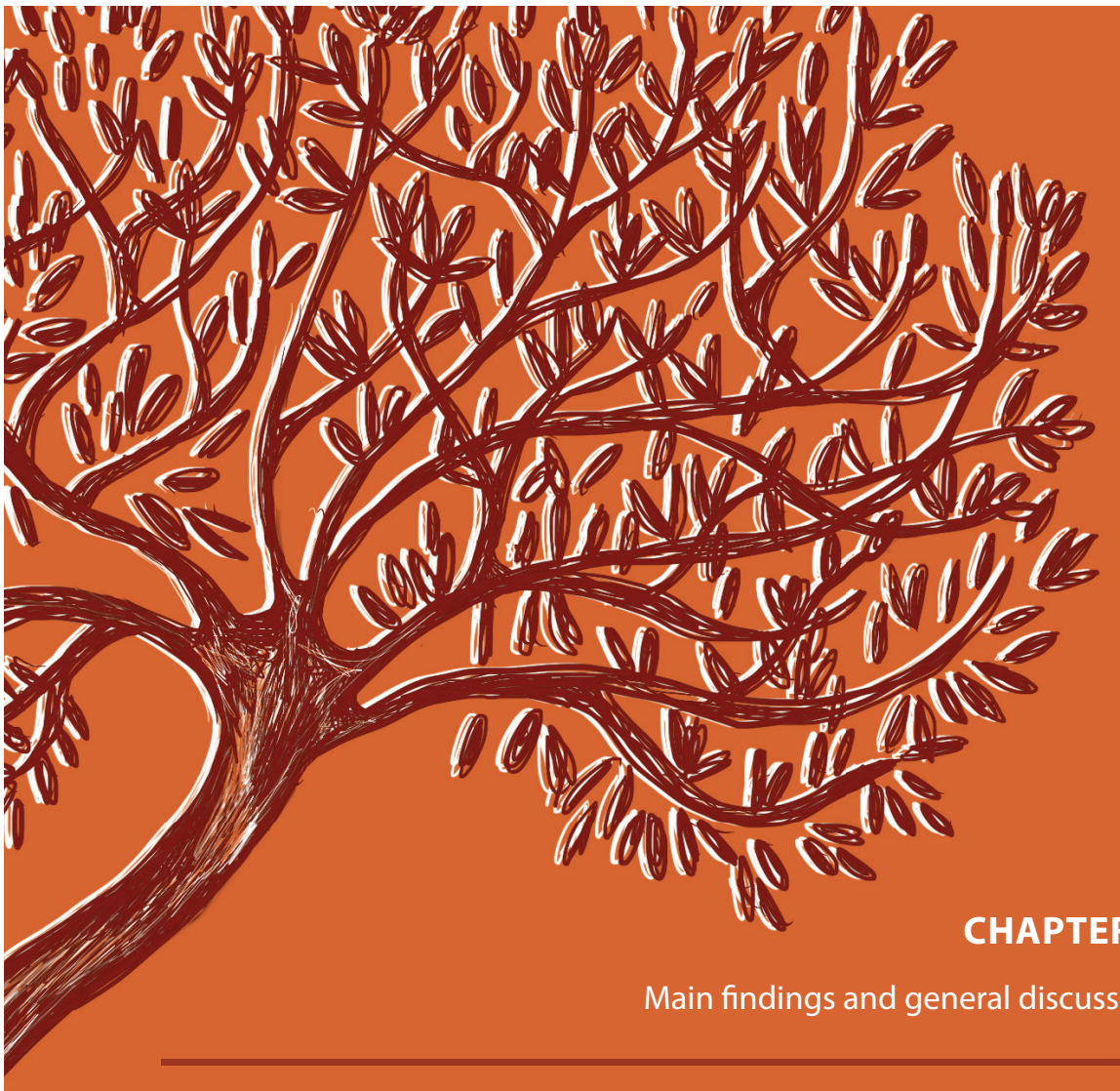
The online version contains supplementary material available at https://doi.org/10.1016/j.jpsychires.2022.09.028

References

1. Kho KH, van Vreeswijk MF, Simpson S, Zwinderman AH. A meta-analysis of electroconvulsive therapy efficacy in depression. *J ECT*. 2003;19(3):139-147.
2. Geddes J, Carney S, Cowen P, et al. Efficacy and safety of electroconvulsive therapy in depressive disorders: A systematic review and meta-analysis. *Lancet*. 2003;361(9360):799-808. doi:10.1016/S0140-6736(03)12705-5
3. Semkovska M, McLoughlin DM. Objective cognitive performance associated with electroconvulsive therapy for depression: A systematic review and meta-analysis. *Biol Psychiatry*. 2010;68(6):568-577. doi:10.1016/j.biopsych.2010.06.009
4. Vasavada MM, Leaver AM, Njau S, et al. Short- and Long-term Cognitive Outcomes in Patients with Major Depression Treated with Electroconvulsive Therapy. *J ECT*. 2017;33(4):278-285. doi:10.1097/YCT.0000000000000426
5. Dybedal GS, Tanum L, Sundet K, Gaarden TL, Bjølseth TM. Cognitive side-effects of electroconvulsive therapy in elderly depressed patients. *Clin Neuropsychol*. 2014;28(7):1071-1090. doi:10.1080/13854046.2014.958536
6. Obbels J, Verwijk E, Vansteelandt K, et al. Long-term neurocognitive functioning after electroconvulsive therapy in patients with late-life depression. *Acta Psychiatr Scand*. 2018;138(3):223-231. doi:10.1111/acps.12942
7. Merikangas K, Jin R, He J, et al. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Arch Gen Psychiatry*. 2011;68(3):241-251. doi:10.1001/ARCHGENPSYCHIATRY.2011.12
8. Ingram A, Saling MM, Schweitzer I. Cognitive side effects of brief pulse electroconvulsive therapy: A review. *J ECT*. 2008;24(1):3-9. doi:10.1097/YCT.0b013e31815ef24a
9. Verwijk E, Comijs HC, Kok RM, Spaans H-P, Stek ML, Scherder EJA. Neurocognitive effects after brief pulse and ultrabrief pulse unilateral electroconvulsive therapy for major depression: A review. 2012. doi:10.1016/j.jad.2012.02.024
10. Jung T, Wickrama KAS. An Introduction to Latent Class Growth Analysis and Growth Mixture Modeling. *Soc Personal Psychol Compass*. 2008;2(1):302-317. doi:10.1111/j.1751-9004.2007.00054.x
11. Proust-Lima C, Philipps V, Liquet B. Estimation of extended mixed models using latent classes and latent processes: The R package lcmm. *J Stat Softw*. 2017;78. doi:10.18637/jss.v078.i02
12. Beltz AM, Wright AGC, Sprague BN, Molenaar PCM. Bridging the Nomothetic and Idiographic Approaches to the Analysis of Clinical Data. *Assessment*. 2016;23(4):447-458. doi:10.1177/1073191116648209
13. Consoloni J Lou, M'Bailara K, Perchec C, et al. Trajectories of medication adherence in patients with Bipolar Disorder along 2 years-follow-up. *J Affect Disord*. 2021;282(December 2020):812-819. doi:10.1016/j.jad.2020.12.192

14. Sha T, Cheng W, Yan Y. Prospective association between sleeprelated factors and the trajectories of cognitive performance in the elderly Chinese population across a 5-year period cohort study. *PLoS One*. 2019;14(9):1-17. doi:10.1371/journal.pone.0222192
15. Sheehan D V., Lecrubier Y, Sheehan KH, et al. 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. *J Clin Psychiatry*. 1998;59(SUPPL. 20):22-33. doi:10.1016/S0924-9338(99)80239-9
16. Hamilton M. A RATING SCALE FOR DEPRESSION. *J Neurol Neurosurg Psychiatry*. 1960;23(1):56-62. doi:10.1136/JNNP.23.1.56
17. Trajković G, Starčević V, Latas M, et al. Reliability of the Hamilton Rating Scale for Depression: a meta-analysis over a period of 49 years. *Psychiatry Res*. 2011;189(1):1-9. doi:10.1016/j.psychres.2010.12.007
18. van Diermen L, Hebbrecht K, Schrijvers D, Sabbe BCG, Fransen E, Birkenhäger TK. The Maudsley Staging Method as predictor of electroconvulsive therapy effectiveness in depression. *Acta Psychiatr Scand*. 2018;138(6). doi:10.1111/acps.12962
19. van Diermen L, Vanmarcke S, Walther S, et al. Can psychomotor disturbance predict ect outcome in depression? *J Psychiatr Res*. 2019;117(May):122-128. doi:10.1016/j.jpsychires.2019.07.009
20. van Diermen L, Versyck P, van den Ameele S, et al. Performance of the Psychotic Depression Assessment Scale as a Predictor of ECT Outcome. *J ECT*. 2019;00(00):1. doi:10.1097/YCT.0000000000000610
21. Nasreddine ZS, Phillips NA, BÃ©dirian V, et al. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *J Am Geriatr Soc*. 2005;53(4):695-699. doi:10.1111/j.1532-5415.2005.53221.x
22. Hebbrecht K, Giltay EJ, Birkenhäger TK, et al. Cognitive change after electroconvulsive therapy in mood disorders measured with the Montreal Cognitive Assessment. *Acta Psychiatr Scand*. 2020;142(5):413-422. doi:10.1111/acps.13231
23. Jaeger J. Digit symbol substitution test. *J Clin Psychopharmacol*. 2018;38(5):513-519. doi:10.1097/JCP.0000000000000941
24. Benedict RHB, Schretlen D, Groninger L, Brandt J. Hopkins Verbal Learning Test – Revised: Normative Data and Analysis of Inter-Form and Test-Retest Reliability. *Clin Neuropsychol*. 1998;12(1):43-55. doi:10.1076/clin.12.1.43.1726
25. Kopelman MD, Wilson BA, Baddeley AD. The autobiographical memory interview: a new assessment of autobiographical and personal semantic memory in amnesic patients. *J Clin Exp Neuropsychol Off J Int Neuropsychol Soc*. 1989;11(5):724-744. doi:10.1080/01688638908400928
26. Fraser LM, O'Carroll RE, Ebmeier KP. The effect of electroconvulsive therapy on autobiographical memory: A systematic review. *J ECT*. 2008;24(1):10-17. doi:10.1097/YCT.0B013E3181616C26
27. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134(4):382-389. doi:10.1192/bjp.134.4.382
28. *National Institute for Clinical Excellence (NICE). 2003. Guidance on the Use of Electroconvulsive Therapy. National Inst. Clin. Excellence*.

29. Nederlandse Vereniging voor Psychiatrie (NVVP). 2010. Richtlijn elektroconvulsietherapie. De Tijdstroom.
30. Hinton-Bayre AD. Deriving reliable change statistics from test-retest normative data: Comparison of models and mathematical expressions. *Arch Clin Neuropsychol*. 2010;25(3):244-256. doi:10.1093/arclin/acq008
31. Jacobson NS, Truax P. Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol*. 1991;59(1):12-19. doi:10.1037//0022-006x.59.1.12
32. Duff K. Current topics in science and practice evidence-based indicators of neuropsychological change in the individual patient: Relevant concepts and methods. *Arch Clin Neuropsychol*. 2012;27(3):248-261. doi:10.1093/arclin/acr120
33. Duff K. One-Week Practice Effects in Older Adults: Tools for Assessing Cognitive Change. *Clin Neuropsychol*. 2014;28(5):714-725. doi:10.1080/13854046.2014.920923
34. Pereira DR, Costa P, Cerqueira JJ. Repeated assessment and practice effects of the written Symbol Digit Modalities Test using a short inter-test interval. *Arch Clin Neuropsychol*. 2015;30(5):424-434. doi:10.1093/arclin/acv028
35. Nuninga JO, Claessens TFI, Somers M, et al. Immediate and long-term effects of bilateral electroconvulsive therapy on cognitive functioning in patients with a depressive disorder. *J Affect Disord*. 2018;238:659-665. doi:10.1016/j.jad.2018.06.040
36. Rose D, Fleischmann P, Wykes T, Leese M, Bindman J. Patients' perspectives on electroconvulsive therapy: systematic review. *BMJ*. 2003;326(7403):1363-1365. doi:10.1136/BMJ.326.7403.1363
37. Donahue AB. Electroconvulsive therapy and memory loss: a personal journey. *J ECT*. 2000;16(2):133-143. http://www.ncbi.nlm.nih.gov/pubmed/10868323. Accessed August 21, 2017.
38. Squire L, Slater P. Electroconvulsive therapy and complaints of memory dysfunction: a prospective three-year follow-up study. *Br J Psychiatry*. 1983;142(1):1-8. doi:10.1192/BJP.142.1.1
39. Verwijk E. Neurocognitive performance in electroconvulsive therapy: To Lose Or Not To Lose? 2015.
40. Roca M, Monzón S, Vives M, et al. Cognitive function after clinical remission in patients with melancholic and non-melancholic depression: a 6 month follow-up study. *J Affect Disord*. 2015;171:85-92. doi:10.1016/J.JAD.2014.09.018
41. Kirov GG, Owen L, Ballard H, et al. Evaluation of cumulative cognitive deficits from electroconvulsive therapy. *Br J Psychiatry*. 2016;208(3):266-270. doi:10.1192/bjp.bp.114.158261
42. Lisanby SH, McClintock SM, McCall W V., et al. Longitudinal Neurocognitive Effects of Combined Electroconvulsive Therapy (ECT) and Pharmacotherapy in Major Depressive Disorder in Older Adults: Phase 2 of the PRIDE Study. *Am J Geriatr Psychiatry*. 2022;30(1):15-28. doi:10.1016/J.JAGP.2021.04.006



CHAPTER 9

Main findings and general discussion

This chapter of the thesis discusses the main findings of our investigations into the etiopathogenesis and course and of mood disorders. Although researchers have accrued important knowledge in the last decades, traditional approaches to mood disorder research have some limitations. Two of these issues are that inferences based on group-level data do not generalize well to individual patients, and that the dynamic nature of depression symptoms is often not taken into account. In the research reported on in this dissertation, we have used traditional approaches but also expounded on approaches that overcome these two issues.

9.1. Traditional approaches

In the first part of the thesis, we investigated the symptom profiles, course and inflammatory biomarkers of Major Depressive Disorder (MDD) and/or Bipolar Disorder (BD) using three more conventional research approaches. Specifically, we studied the clinical and biological variables in relation to syndrome-based entities such as MDD or BD by analyzing group-level averages of outcome variables (e.g., mean scores of depression rating scales or mean peripheral blood levels of inflammatory markers). Changes in average severity scores of mood items or biological markers over time were investigated without taking their mutual dynamic interactions into account (as opposed to the dynamic approach in Chapter 9.2).

9.1.1. Symptom profile and course of depressive episode in patients with MDD versus BD (Chapter 2)

Bipolar depression, especially in the absence of a history of a (hypo)manic episode, is often misdiagnosed as a unipolar depression, leading to suboptimal treatment and prognosis [1–4]. Antidepressant monotherapy in BD is discouraged due to its low efficacy and risk of switch to (hypo)mania or acceleration of mood cycles [5–7]. In Chapter 2, we studied the presence of clinically discriminatory features of a Major Depressive Episode (MDE) in a naturalistic sample of inpatients with MDD or BD. The symptom profiles and the mean time to response and remission were compared among patients diagnosed with MDD (n = 224) and those with a BD diagnosis (n = 52) being treated for a current MDE using repeated short-term (i.e. 2-weekly) routine outcome measurements (ROM).

Our analyses revealed only small differences in average symptom scores in MDD and BD patients on the group-level, with inpatients with MDD scoring higher on the HRSD-17 item weight loss. The obtained symptom differences between MDD and BD were consistent with available literature reviews [8,9]. Yet, for individual clinical patients, these group-level findings showed not enough discriminatory power to be of clinical value.

With regard to the prognosis, we found that inpatients with BD showed a higher likelihood of response than inpatients with MDD after 28 days of hospitalization. This finding was largely consistent with previous studies [10,11], but was inconsistent with findings from the study of Kessing e.a. [12] that reported no difference in the rate of recovery between the two disorders.

It must be acknowledged that the HRSD-17 provides only a limited assessment of depressive symptoms. Previous studies have also shown more atypical depression symptoms (e.g., hypersomnia, hyperphagia, leaden paralysis) [13–19], psychomotor retardation [13,16,20] and psychotic symptoms [13,17–19] in depressed BD patients than in those with unipolar depression, but these are not (or not adequately) being assessed by the HRSD-17. The HRSD-17 was originally developed for inpatients with severe depression and relies heavily on clinical, observable signs such as weight loss and slowing of speech rather than self-reported more subjective symptoms [21]. Also, we made no distinction between type-I or type-II BD although there is some evidence to suggest different depressive symptom profiles for the two types (e.g., both show atypical features, but no consistent finding of more psychomotor retardation in BD type II).

Concluding, although there appear to be no general and consistent findings on a symptom or course profile that distinguishes a patient with bipolar depression from a patient with unipolar depression, some clinical features may provide diagnostic clues aiding their distinction.

9.1.2. Tryptophan catabolites as a biomarker for BD (Chapters 3 and 4)

Accumulating evidence points to the possible pathophysiological role of chronic low-grade inflammation in mood disorders [22–25]. The kynurenine pathway of tryptophan degradation has been put forward as the missing link explaining how inflammation could mediate the development of mood [26–28] and cognitive symptoms in BD [27]. A better understanding of the etiopathogenetic role of TRYCATs in BD could pave the way for more effective treatments. To contribute to this quest, in Chapter 3 and 4, we investigated the potential involvement of a dysregulation of the TRYCAT pathway in BD and its associated cognitive dysfunction. First, we performed a systematic review and meta-analysis of the available literature on TRYCAT alterations in the cerebrospinal fluid (CSF) or peripheral blood of individuals with BD versus healthy controls (Chapter 3), then we explored the associations between TRYCAT levels and cognitive measures in a study sample of patients with BD (35 depressed and 32 (hypo)manic patients) compared with healthy controls (n = 32) (Chapter 4).

Our meta-analysis (Chapter 3) revealed three main findings. First, peripheral levels of TRP, KYN and KA were lower in patients with BD than they were in the controls. Second,

peripheral levels of 3-HK and QA were not significantly between the two groups although we expected these to be increased under inflammatory conditions [27,29,30]. Third, TRP and KYN alterations showed no differential patterns across mood states. The original research study in Chapter 4 showed stable low levels of KA in BD patients, compared to healthy controls, and these were significantly associated with lower scores for overall cognitive functioning, but only in the depressed group.

The publication of our meta-analysis coincided with the publication of two other meta-analyses [31,32]. Opposed to our study, however, Marx et al. [31] did find some differences in depressed, manic and euthymic patients with BD. Specifically, they found lower KYN and QA concentrations in the depressed BD group but not in the manic and euthymic groups. This was probably due to differences in included studies (among which also unpublished data) compared to our study. The lower peripheral TRP levels in mood disorders support the serotonin depletion hypothesis [33–35], but they do not entirely reconcile with the theoretically proposed hypothesis of an increased TRP breakdown towards the neurotoxic branch (i.e. lower KYN and KA levels; higher levels of 3-HK and QA). A possible explanation may be the upregulation of the microglial branch resulting in reduced KA levels [36]. Furthermore, compared to other TRYCAT, studies examining QA and 3-HK levels are scarce and more subtle alterations in their respective concentrations may be more difficult to identify due to the assay's limited detection rate [37]. The association between low KA levels and lower scores for overall cognitive functioning could reflect a critically low concentration below which the neuroprotective potential of KA suddenly decreases. However, it should be noted that not all cognitive test scores showed a negative association with KA levels and the more cognitively impaired manic BD group did not show this association at all. A recently published meta-analysis by Morrens et. al. [38] also found only weak associations between a wide range of blood-based immune markers (including TRYCAT) and cognitive measures.

It is important to note that most of our findings relate to peripheral blood levels rather than cerebrospinal fluid (CSF) levels. Where current evidence suggests that TRP, KYN and 3-HK cross the blood-brain barrier [39], a recent literature study reported that only peripheral levels of KYN and 3-HK appear to reflect their central levels [40]. This could explain the discrepancy in the various findings, where a previous meta-analysis found the KA levels in CSF to be elevated [41], while our and other meta-analyses observed reduced concentrations of peripheral KA [31,32,42]. Nevertheless, the low peripheral KA levels in BD could still be a valuable biomarker as they have been associated with central nervous system processes/ biological features of BD (e.g., reduced brain connectivity) in previous studies.

Together, we found evidence for the involvement of a dysregulation of the TRYCAT pathway in BD. However, its exact pathophysiological mechanism remains to be

elucidated. Furthermore, effect-sizes were small rendering the clinical relevance of our, and previous, findings questionable. It is highly probable that TRYCAT alterations only play a partial etiopathogenetic role in a subset of patients with BD.

9.2. Idiographic approach to symptom dynamics

In the studies presented in the second part of this thesis, we aimed to gain more insight into the course of depression in MDD and BD by taking upon an alternative approach. Specifically, we focused on individual symptoms rather than whole syndromes and aimed to capture their dynamic behaviour and interplay with other symptoms over time. These dynamic symptom changes were investigated at both the individual-level and the group-level.

9.2.1. Undirected dynamics of Hamilton Rating Scale for Depression-17 symptoms (Chapter 5)

Mood disorders are heterogeneous and multifactorial disorders. These disorders do not result from a common cause but can better be modelled as a complex dynamic system in which the correlation between symptoms is regarded as resulting from interactions of components within the system (instead of stemming from some common cause). The system is dynamic in that the symptom-symptom relations can evolve over time within an individual. Patient-specific change profiles of depression symptoms over time are thus likely to yield a more detailed picture of processes at play in MDE that may have clinical relevance [43–45]. Also, symptoms showing similar change over time may reflect a shared etiopathogenesis or a direct causal relationship [46]. Changes in sleep duration and weight gain, for instance, may cluster differently in subgroups of patients, while in patients with an atypical depression their coordinated waning and waxing may be indicative of a causal relationship or a shared underlying causal mechanism.

In Chapter 5, we aimed to estimate symptom dimensions (i.e., groups of symptoms with similar dynamics over time), using undirected DTW analysis and hierarchical clustering of repeated Hamilton Depression Rating Scale (HRSD-17) measurements. These analyses were performed in 255 patients with MDE (either MDD or BD), both at the level of the individual patient (i.e., symptom dimensions within each patient) and the group-level (i.e., dimensions representing systematic patterns of symptom dynamics across patients). Additionally, we investigated the relationship between network density and reaching remission by examining the density of the overall dynamic symptom network in patients reaching remission and non-remitters.

The patient-specific dimensions varied considerably across patients. Group-level analyses revealed five HRSD-17 symptom dimensions: 'Core Depressive Symptoms', 'Social Withdrawal and Lethargy', 'Disturbed Self-definition' and 'Auto-aggression'. As these five symptom dimensions show different trajectories over time, only using sum scores of the total HRSD-17 may mask important clinical information of symptom courses over time. The items from each of five dimensions could be added to yield five subscores, which may be used in clinical practice to follow individual patients over time, according to their trajectories along any of these five dimensions.

We also found that patients failing to reach remission showed a more loosely connected dynamic symptom network compared to remitters. This could indicate that in non-remitters, symptoms behave more independently from each other over time which, in turn, may lower the chance of a 'domino tile effect' or reinforcing feedback loops within the network of symptoms in response to external triggers and hereby lowers the chance of remission to treatment. This finding is in accordance with literature from the Complex Dynamics System Theory: systems with a dense connectivity between the elements can change more abruptly (either towards a healthy or a depressed state) often in response to external perturbations such as stress or treatment (i.e., so called tipping points or phase transitions) [47].

Together, individual patients with MDE showed large variability in their dynamic symptom networks. The combination of individual-level and group-level DTW analyses provides a promising method to gain insight into the complex idiosyncratic nature of depression symptom dynamics.

9.2.2. Directed symptom dynamics of the Beck Depression Inventory-II symptoms (Chapter 6)

A crucial step forward is to gain insight in which symptom changes precede changes in other symptoms. Granger causality determines whether one time series can help in forecasting another, using prior values of the prior time series. Such understanding may yield insights for improved prediction and ultimately for targets of treatment intervention.

Therefore, in Chapter 6, we performed both directed DTW analyses (assessing the directionality of symptom change) in addition to undirected DTW analyses (i.e., investigating similarity in dynamic behavior over time) on the bi-weekly BDI-II measurements. Patients with MDE (either MDD or BD) and a minimum of four consecutive BDI-II measurements were included in the analyses, which resulted in a study sample of 166 patients. Both DTW analyses were first done at the level of the individual patient and the data were subsequently aggregated to the group level.

Four undirected BDI-II symptom dimensions, consisting of symptoms with similar dynamics over time, could be distinguished: 'Core Depressive Symptoms', 'Social Withdrawal and Lethargy', 'Disturbed Self-definition' and 'Auto-aggression'. Group-level directed analyses showed that the symptoms self-criticism and loss of interest had a significantly higher instrength score, supporting the idea that these symptoms are highly influenced by other symptoms. The symptoms worthlessness, pessimism, tiredness and sadness, having the highest outstrength, presumably more often preceded changes in other symptoms.

Concluding, directed DTW analysis may help to deepen our understanding of MDE through a complex dynamic system lens. Symptoms with high outstrength could be potentially useful treatment targets, however this should be further investigated using experimental study designs.

9.3. A combined individual- and group-level approach into cognitive side effects of ECT (Chapters 7 and 8)

Apart from mood symptoms, cognitive symptoms are central features in mood disorders that can be highly disabling [48–51]. Even though the interindividual variability of cognitive effects of ECT is common knowledge [52–54], the cognitive effects of ECT are mainly investigated using aggregate test scores, whereby cognitive functioning is treated as a homogeneous phenomenon across patients and time.

In Chapter 7 and 8, we therefore attempted to get a more nuanced view of ECT-related cognitive functioning in patients with MDE by combining three analytic approaches. Specifically, cognitive functioning was investigated in ECT-treated patients with MDE using a test battery assessing global cognitive functioning, processing speed, verbal episodic memory, and retrograde memory. Changes in cognitive functioning over time were analyzed at the group-level using linear mixed models. Individual-specific changes in cognitive function were examined using Reliable Change Indices (RCI). Lastly, the presence of subgroups patients with similar cognitive trajectories over time was investigated using Latent Class Growth Analysis (LCGA).

Our results showed no lasting ECT-induced cognitive deficits at the group level, while the individual-level analyses revealed substantial heterogeneity in cognitive performance during and after ECT. More specifically, there was a transient cognitive deterioration in the group-level averaged test scores of processing speed, verbal episodic (learning and delayed recall) memory and retrograde (personal semantic) memory during and immediately after ECT completion, with all average scores having returned to baseline values at the 3- and 6-month follow-ups. Global cognitive functioning showed no significant group changes

compared to baseline. At the individual level, there was substantial heterogeneity with subgroups of patients showing either deterioration, improvement or no change compared to baseline at short-term or long-term follow-up.

Based on their longitudinal trajectories of cognitive performance, patients could be categorized into three classes. These classes had a fairly parallel courses, being characterized by an initial decline followed by a recovery to baseline levels; the groups primarily differed with respect to the average baseline cognitive functioning. Univariate analyses showed that class membership was mainly determined by the individual's baseline Montreal Cognitive Assessment (MoCA) scores and the presence of melancholic features at baseline.

Concluding, no lasting ECT-induced cognitive deficits were found at the group level, while patient-specific analyses revealed substantial inter-individual variability. Cognitive changes over time appeared to be fairly similar among patients and were mainly determined by differences in baseline variables (MoCA and/or presence of melancholic features), not by improvement in mood or ECT parameters.

9.4. General discussion

The studies in Part 1 (Chapter 2 to 4) are founded on widely used research paradigms, in which mood disorders are seen as discretely delineated categories and mean sum-scores of outcome variables are compared on the group-level. Although these approaches have resulted in many insights that form the fundaments of our current knowledge in psychiatry, research progress is rather slow, and some fundamental issues may lie at the base thereof. Furthermore, traditional approaches do not take the dynamic nature of mood disorders (and its constituent variables such as psychological and biological elements) into account, while many psychiatric disorders are thought of to be the result of such dynamic interactions [21,55].

The lack of general and consistent findings on clinical discriminatory features for a depression in either MDD or BD in our (Chapter 2) and previous studies adds to a long-lasting debate whether MDD and BD should rather be represented as an affective disorders continuum [56–59] (instead of the traditional unipolar-bipolar categorical dichotomy). Several authors have stated the importance of adopting a dimensional approach for the study of mood disorders. It is hypothesized that dimensional measures are strongly linked to the pathophysiological processes in psychiatric disorders compared to the traditional DSM categories.

The use of categorical (DSM) diagnoses may also have contributed to the small effect sizes on the associations between TRYCAT metabolite alterations and BD in ours and previous meta-analyses. Alterations in TRYCAT levels have been shown to transcend traditional diagnostic boundaries as partly overlapping TRYCAT alterations have been found in MDD, BD and schizophrenia [31]. As stated by previous authors [60–62], it is plausible that the studies included in our meta-analysis have adopted a too reductionist view by studying isolated biological variables and syndrome-based groups of mood disorders. In line with the RDoC's focus, future research could benefit from studying the study the relationship between inflammatory (and TRYCAT) alterations in relation to symptom subsets, rather than in relation to a unified syndrome (e.g., MDD or BD), as distinct etiopathogenetic processes may have a unique relationship to symptoms subset. For example, Van Eeden e.a. [63] found basal and lipopolysaccharide-induced inflammatory markers to be more strongly linked to sickness behavior symptoms (e.g., anhedonia, anorexia, low concentration, low energy, loss of libido, psychomotor slowness, irritability, and malaise)

Part 2 of this thesis provided an alternative investigation of the course of MDE by studying individual symptom dynamics over time, both at the individual- and the group level. Our findings highlighted the idiosyncratic behavior of depression symptoms over time and hereby support the concern about the parsimonious use of sum scores of depression rating scales as an outcome measure in mood disorder research [64,65]. Although these sum scores represent easily interpretable values, they entirely disregard the heterogeneity of depression symptoms between patients. Furthermore, analyzing the change in sum scores over time to investigate the course of MDE over time does injustice to the complexity of dynamic symptom interactions and hereby result in a loss of potentially relevant clinical information. DTW analyses can help to construct a graphical representation of the dynamic interplay among symptoms both at the level of the individual patient and the group-level. Although these had fixed structures it is dynamic in the sense that the nodes take on severity scores that change over time. As these changes are correlated, the behavior of the system becomes dynamic. We have studied the directed associations which may represent positive feedback dynamics. The dynamic nature also rests on negative feedback loops, which helps symptom networks to become resistant to small-scale perturbation, including stressors and noise. This resistance could be embedded in the symptom organization. As we solely studies patients, and no healthy controls, it is possible that our findings represent the more unstable symptom organizations of those with a mood disorder.

The exploratory DTW analyses were able to visualize the idiosyncratic nature of symptom dynamics and the dynamic symptom-symptom interactions over time. Our work should be considered as a stepping stone towards further research on symptom dynamics in mood disorders. A first suggestion for future research is to evaluate the replicability and

robustness of our findings and explore the impact of changes in default setting of the DTW method. Second, future studies should aim to compare DTW with other existing analytic methods to study dynamic symptom relations (e.g., lagged multilevel vector autoregressive models) [66]. Third, there is an ongoing debate as to which symptoms exactly constitute depression. Future studies into symptom dynamics are then advised to investigate as wide an array of symptoms as possible, for instance by using the 44-item Symptoms of Depression Questionnaire [67] or a combination of (self-report and clinician-rated) assessment instruments. These scales should ideally consist of multiple questions for each symptom to control for measurement error and reduce the effects of random noise [46]. The use of the HRSD-17 and the BDI-II in our studies, with their three-to-five point Likert scales that can be considered rather crude, may have resulted in low variability and precision. Assessment scales should be selected and applied with the sole goal of evaluating symptoms and not to gauge a supposedly underlying unidimensional construct [68]. Fifth, the temporal behavior of symptoms should be further studied in order to solve the open question on which time intervals symptoms should be assessed. Our intervals of 2 weeks may have overlooked faster-occurring symptom changes. Finally, research should not only focus on depression symptoms but should widen its focus and study multiple components across biological, psychological, and social levels of analysis and the complex web of interactions that is formed between these components [69]. For example, studies in the field of clinical psychology have focused on relations between different components that operate on different time-scales (i.e., the micro-, meso- macro-level) [70,71]. Emotions and symptoms can be defined as transient subjective states that change in response to internal or external stimuli and lasting within a range of hours to days (micro-level). Mood episodes are internal subjective states typically lasting for weeks to months (meso-level) and personality traits (e.g., neuroticism) represent more persisting traits over time and can be seen as "set-points" around which the micro- and meso-level components fluctuate in response to life events [72]. Studies that aim to capture these various dynamic processes, across multiple time scales, and changes therein over time shed light on the emergence of psychopathology on the micro, meso and macro-level [72]. Modern technologies like smart-phone apps can be used to collect such data.

We did not investigate the effects of specific clinical interventions in our DTW studies. We therefore need to be cautious with formulating clinical implications. We will however discuss some potential clinical implications while integrating them with future research goals [46,73]. First, future research should examine whether these complex symptom dynamics have added clinical value over mean symptom scores in the prediction of the course and outcome of MDE or prediction of overall wellbeing [74]. Second, the role of directed DTW analyses for intervention strategies should be further investigated. Directed DTW analyses allow the identification of symptoms with high outstrength that represent highly connected symptoms which presumably trigger downstream activation of other

symptoms. Our directed DTW analyses revealed that the symptom worthlessness had a high outstrength indicating that changes in feelings of worthlessness were often followed in time by similar changes in other symptoms such as self-criticism, loss of energy and insomnia. Future studies could for example test the hypothesis whether cognitive behavior intervention targeted at DTW symptoms with high outstrength leads to a more rapid alleviation of depression symptoms.

The last part of this thesis includes two ECT studies that combined group- and individual-level analyses to provide a more nuanced insight into the cognitive side effects of ECT. Our group-level findings add to the reassuring evidence that cognitive side effects of ECT are predominantly transient [75]. However, these reassuring results on the group level need to be interpreted with caution. First, it should be emphasized that pre-ECT levels most likely reflect impaired cognitive function as the depression itself has a negative influence on cognitive function [49]. So, poor cognitive performance scores at follow-up that are comparable to baseline could be considered to reflect persistent impairment, not compared to pre-ECT levels but compared to predepression performance. Several hypotheses have been proposed to explain this persistence. One hypothesis states that ECT does have longer lasting cognitive side effects that are masked by the positive influence of mood improvements on cognition. The so-called "scar hypothesis" poses that depressive episodes might cause residual cognitive impairments ("cognitive scars") rendering patients increasingly cognitive vulnerable for subsequent episodes [76]. A third hypothesis states that some types of cognitive dysfunction could represent trait characteristics in patients with an increased vulnerability for developing mood disorders [77]. It should also be noted that group-level averages overlook important interindividual differences in cognitive effects of ECT. Our studies corroborate previous findings that small groups of patients show persisting cognitive effects (and, similarly, small groups of patients show cognitive improvement) after treatment with ECT. The RCI analyses provide clinically relevant information that supplement group-level findings.

Future research should be aimed at further elucidating the observed interindividual variability in cognitive performance in patients treated with ECT. Specifically, the relationship between ECT-related and mood-related cognitive changes in patients with MDE is poorly understood and needs further investigation. Research should therefore aim to identify predictors for ECT-related cognitive impairment, and their interdependence, within a comprehensive framework. Mc Clintock [78] proposed such a conceptual model in which several moderating (e.g., demographic and clinical factors) and mediating factors (e.g., ECT parameters and neurophysiological changes) are incorporated to examine the cognitive changes during an ECT care pathway. Another topic for further study is the variation in cognitive assessment strategy between different studies. Future research should answer the question which cognitive test battery shows adequate sensitivity to

ECT-related cognitive changes and should be uniformly used in research (and clinical practice).

The inter-individual variability in cognitive performance during an ECT care pathway that was observed in our two studies highlights the importance of routine cognitive measurements in clinical ECT practice. These repeated assessments enable the early recognition of the minority group of patients with a persisting cognitive impairment after ECT, with a potentially great impact on psychosocial functioning. Further, it allows a fine-tuning of the ECT parameters accordingly with a potential to improve clinical ECT care. Lastly, the cognitive test outcomes could be used as a feedback tool to discuss cognitive side effects with patients with a possible reduction in early discontinuation rates.

9.5 Considerations

The studies presented in this thesis have some inherent methodological limitations that need to be discussed. First, the results presented in this thesis are based on data from three study projects with an observational design which hampers explicit causal inference and clinical translation of the findings. Second, the PROTECT study (Chapter 7, 8) did not include a control sample, which makes it impossible to state whether the described cognitive changes are specific for ECT treatment. Third, the HERCULES ROM (Chapter 2, 5, 6) study included only inpatients from a single tertiary psychiatric hospital which may have negatively influenced the generalizability. Fourth, in the studies discussed in Chapter 2, 3, 7 and 8 we adjusted for key potential confounding variables. However, not all were measured so that residual confounding is possible (e.g., we could not adjust for psychotropic drug use). Lastly, the study of individual depression symptoms dynamics (Chapter 5 and 6) was based on HRSD-17 and BDI-II items. However, it should be noted that single items are more strongly influenced by random and measurement error than sum scores [46]. As already stated above, future studies could more accurately assess symptoms by using multi-item measures.

References

1. Li CT, Bai YM, Huang YL, et al. Association between antidepressant resistance in unipolar depression and subsequent bipolar disorder: cohort study. *Br J Psychiatry*. 2012;200(1):45-51. doi:10.1192/BJP.BP.110.086983
2. Hu C, Xiang YT, Ungvari GS, et al. Undiagnosed bipolar disorder in patients treated for major depression in China. *J Affect Disord*. 2012;140(2):181-186. doi:10.1016/J.JAD.2012.02.014
3. Altamura AC, Dell'Osso B, Berlin HA, Buoli M, Bassetti R, Mundo E. Duration of untreated illness and suicide in bipolar disorder: a naturalistic study. *Eur Arch Psychiatry Clin Neurosci*. 2010;260(5):385-391. doi:10.1007/S00406-009-0085-2
4. Post RM, Leverich GS, Kupka RW, et al. Early-Onset Bipolar Disorder and Treatment Delay Are Risk Factors for Poor Outcome in Adulthood. *J Clin Psychiatry*. 2010;71(7):0-0. doi:10.4088/JCP.08M04994YEL
5. Kendall T, Morriss R, Mayo-Wilson E, Marcus E. Assessment and management of bipolar disorder: summary of updated NICE guidance. *BMJ*. 2014;349(sep25 5):g5673-g5673. doi:10.1136/bmj.g5673
6. Milev R V., Giacobbe P, Kennedy SH, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder. *Can J Psychiatry*. 2016;61(9):561-575. doi:10.1177/0706743716660033
7. Goodwin GM, Haddad PM, Ferrier IN, et al. Evidence-based guidelines for treating bipolar disorder: Revised third edition recommendations from the British Association for Psychopharmacology. *J Psychopharmacol*. 2016;30(6):495-553. doi:10.1177/0269881116636545
8. Mitchell PB, Goodwin GM, Johnson GF, Hirschfeld RM. Diagnostic guidelines for bipolar depression: a probabilistic approach. *Bipolar Disord*. 2008;10(1p2):144-152. doi:10.1111/j.1399-5618.2007.00559.x
9. Angst J, Azorin J-M, Bowden CL, et al. Prevalence and characteristics of undiagnosed bipolar disorders in patients with a major depressive episode: the BRIDGE study. *Arch Gen Psychiatry*. 2011;68(8):791-798. doi:10.1001/archgenpsychiatry.2011.87
10. Winokur G, Coryell W, Keller M, Endicott J, Akiskal H. A prospective follow-up of patients with bipolar and primary unipolar affective disorder. *Arch Gen Psychiatry*. 1993;50(6):457-465. http://www.ncbi.nlm.nih.gov/pubmed/8498880. Accessed December 7, 2018.
11. Angst J, Preisig M. Course of a clinical cohort of unipolar, bipolar and schizoaffective patients. Results of a prospective study from 1959 to 1985. *Schweiz Arch Neurol Psychiatr*. 1995;146(1):5-16. http://www.ncbi.nlm.nih.gov/pubmed/7792568. Accessed December 7, 2018.
12. Kessing L V, Mortensen PB. Recovery from episodes during the course of affective disorder: a case-register study. *Acta Psychiatr Scand*. 1999;100(4):279-287. http://www.ncbi.nlm.nih.gov/pubmed/10510697. Accessed December 10, 2018.
13. Mitchell PB, Wilhelm K, Parker G, Austin MP, Rutgers P, Malhi GS. The clinical features of bipolar depression: a comparison with matched major depressive disorder patients. *J Clin Psychiatry*. 2001;62(3):212-216; quiz 217.

14. Serretti A, Mandelli L, Lattuada E, Cusin C, Smeraldi E. Clinical and demographic features of mood disorder subtypes. *Psychiatry Res*. 2002;112(3):195-210. http://www.ncbi.nlm.nih.gov/pubmed/12450629. Accessed November 10, 2018.
15. Benazzi F. Clinical differences between bipolar II depression and unipolar major depressive disorder: lack of an effect of age. *J Affect Disord*. 2003;75(2):191-195. http://www.ncbi.nlm.nih.gov/pubmed/12798259. Accessed November 10, 2018.
16. Benazzi F. Symptoms of depression as possible markers of bipolar II disorder. *Prog Neuropsychopharmacol Biol Psychiatry*. 2006;30(3):471-477. doi:10.1016/j.pnpbp.2005.11.016
17. Andreasen NC, Grove WM, Coryell WH, Endicott J, Clayton PJ. Bipolar versus unipolar and primary versus secondary affective disorder: which diagnosis takes precedence? *J Affect Disord*. 15(1):69-80. http://www.ncbi.nlm.nih.gov/pubmed/2970495. Accessed November 10, 2018.
18. Guze SB, Woodruff RA, Clayton PJ. The significance of psychotic affective disorders. *Arch Gen Psychiatry*. 1975;32(9):1147-1150. http://www.ncbi.nlm.nih.gov/pubmed/1180665. Accessed November 10, 2018.
19. Coryell W, Endicott J, Maser JD, Keller MB, Leon AC, Akiskal HS. Long-term stability of polarity distinctions in the affective disorders. *Am J Psychiatry*. 1995;152(3):385-390. doi:10.1176/ajp.152.3.385
20. Parker G, Roy K, Wilhelm K, Mitchell P, Hadzi-Pavlovic D. The nature of bipolar depression: implications for the definition of melancholia. *J Affect Disord*. 2000;59(3):217-224. http://www.ncbi.nlm.nih.gov/pubmed/10854638. Accessed November 10, 2018.
21. Fried EI, Flake JK, Robinaugh DJ. Revisiting the theoretical and methodological foundations of depression measurement. *Nat Rev Psychol*. 2022. doi:10.1038/s44159-022-00050-2
22. Miller AH, Maletic V, Raison CL. Inflammation and Its Discontents: The Role of Cytokines in the Pathophysiology of Major Depression. *Biol Psychiatry*. 2009;65(9):732-741. doi:10.1016/J.BIOPSYCH.2008.11.029
23. Van Den Ameele S, Van Diermen L, Staels W, et al. The effect of mood-stabilizing drugs on cytokine levels in bipolar disorder: A systematic review. *J Affect Disord*. 2016;203:364-373. doi:10.1016/j.jad.2016.06.016
24. Haarman BCMB, Riemersma-Van der Lek RF, de Groot JC, et al. Neuroinflammation in bipolar disorder – A [11C]-(R)-PK11195 positron emission tomography study. *Brain Behav Immun*. 2014;40:219-225. doi:10.1016/J.BBI.2014.03.016
25. Wang YL, Han QQ, Gong WQ, et al. Microglial activation mediates chronic mild stress-induced depressive- and anxiety-like behavior in adult rats. *J Neuroinflammation*. 2018;15(1). doi:10.1186/S12974-018-1054-3
26. Dantzer R, O'Connor JC, Lawson MA, Kelley KW. Inflammation-associated depression: from serotonin to kynurenine. *Psychoneuroendocrinology*. 2011;36(3):426-436. doi:10.1016/J.PSYNEUEN.2010.09.012
27. Savitz J. The kynurenine pathway: a finger in every pie. *Mol Psychiatry*. 2020;25(1):131-147. doi:10.1038/s41380-019-0414-4

28. Schwarcz R, Bruno JP, Muchowski PJ, Wu HQ. Kynurenines in the mammalian brain: When physiology meets pathology. *Nat Rev Neurosci*. 2012;13(7):465-477. doi:10.1038/nrn3257
29. Connor TJ, Starr N, O'Sullivan JB, Harkin A. Induction of indolamine 2,3-dioxygenase and kynurenine 3-monooxygenase in rat brain following a systemic inflammatory challenge: a role for IFN-gamma? *Neurosci Lett*. 2008;441(1):29-34. doi:10.1016/J.NEULET.2008.06.007
30. Molteni R, Macchi F, Zecchillo C, et al. Modulation of the inflammatory response in rats chronically treated with the antidepressant agomelatine. *Eur Neuropsychopharmacol*. 2013;23(11):1645-1655. doi:10.1016/J.EURONEURO.2013.03.008
31. Marx W, McGuinness AJ, Rocks T, et al. The kynurenine pathway in major depressive disorder, bipolar disorder, and schizophrenia: a meta-analysis of 101 studies. *Mol Psychiatry*. 2020. doi:10.1038/s41380-020-00951-9
32. Bartoli F, Misiak B, Callovini T, et al. The kynurenine pathway in bipolar disorder: a meta-analysis on the peripheral blood levels of tryptophan and related metabolites. *Mol Psychiatry*. 2020. doi:10.1038/s41380-020-00913-1
33. Rosa-Neto P, Diksic M, Okazawa H, et al. Measurement of brain regional alpha-[11C]methyl-L-tryptophan trapping as a measure of serotonin synthesis in medication-free patients with major depression. *Arch Gen Psychiatry*. 2004;61(6):556-563. doi:10.1001/ARCHPSYC.61.6.556
34. Vaswani M, Linda FK, Ramesh S. Role of selective serotonin reuptake inhibitors in psychiatric disorders: a comprehensive review. *Prog Neuropsychopharmacol Biol Psychiatry*. 2003;27(1):85-102. doi:10.1016/S0278-5846(02)00338-X
35. Arango V, Underwood MD, Mann JJ. Chapter 35 Serotonin brain circuits involved in major depression and suicide. *Prog Brain Res*. 2002;136:443-453. doi:10.1016/S0079-6123(02)36037-0
36. Garrison AM, Parrott JM, Tuñon A, Delgado J, Redus L, O'Connor JC. Kynurenine pathway metabolic balance influences microglia activity: Targeting kynurenine monooxygenase to dampen neuroinflammation. *Psychoneuroendocrinology*. 2018;94:1-10. doi:10.1016/j.psyneuen.2018.04.019
37. Young SN, Anderson GM. Bioanalytical inaccuracy: a threat to the integrity and efficiency of research. *J Psychiatry Neurosci*. 2010;35(1):3-6. doi:10.1503/JPN.090171
38. Morrens M, Overloop C, Coppens V, et al. The relationship between immune and cognitive dysfunction in mood and psychotic disorder: a systematic review and a meta-analysis. *Mol Psychiatry*. 2022;(February). doi:10.1038/s41380-022-01582-y
39. Kindler J, Lim CK, Weickert CS, et al. Dysregulation of kynurenine metabolism is related to proinflammatory cytokines, attention, and prefrontal cortex volume in schizophrenia. *Mol Psychiatry*. 2020;25(11):2860-2872. doi:10.1038/s41380-019-0401-9
40. Skorobogatov K, De Picker L, Verkerk R, et al. Brain Versus Blood: A Systematic Review on the Concordance Between Peripheral and Central Kynurenine Pathway Measures in Psychiatric Disorders. *Front Immunol*. 2021;12(September). doi:10.3389/fimmu.2021.716980

41. Wang AK, Miller BJ. Meta-analysis of Cerebrospinal Fluid Cytokine and Tryptophan Catabolite Alterations in Psychiatric Patients: Comparisons Between Schizophrenia, Bipolar Disorder, and Depression. *Schizophr Bull*. 2018;44(1):75-83. doi:10.1093/schbul/sbx035
42. Hebbrecht K, Skorobogatov K, Giltay EJ, Coppens V, De Picker L, Morrens M. Tryptophan Catabolites in Bipolar Disorder: A Meta-Analysis. *Front Immunol*. 2021;12(May):1-10. doi:10.3389/fimmu.2021.667179
43. Wichers M. The dynamic nature of depression: A new micro-level perspective of mental disorder that meets current challenges. *Psychol Med*. 2014;44(7):1349-1360. doi:10.1017/S0033291713001979
44. De Vos S, Wardenaar KJ, Bos EH, Wit EC, Bouwmans MEJ, De Jonge P. An investigation of emotion dynamics in major depressive disorder patients and healthy persons using sparse longitudinal networks. *PLoS One*. 2017;12(6). doi:10.1371/JOURNAL.PONE.0178586
45. Hamaker EL, Wichers M. No Time Like the Present: Discovering the Hidden Dynamics in Intensive Longitudinal Data. *Curr Dir Psychol Sci*. 2017;26(1):10-15. doi:10.1177/0963721416666518
46. Fried EI, Cramer AOJ. Moving Forward: Challenges and Directions for Psychopathological Network Theory and Methodology. *Perspect Psychol Sci*. 2017;12(6):999-1020. doi:10.1177/1745691617705892
47. Van De Leemput IA, Wichers M, Cramer AOJ, et al. Critical slowing down as early warning for the onset and termination of depression. *Proc Natl Acad Sci U S A*. 2014;111(1):87-92. doi:10.1073/pnas.1312114110
48. McIntyre RS, Cha DS, Soczynska JK, et al. Cognitive deficits and functional outcomes in major depressive disorder: determinants, substrates, and treatment interventions. *Depress Anxiety*. 2013;30(6):515-527. doi:10.1002/DA.22063
49. Rock PL, Roiser JP, Riedel WJ, Blackwell AD. Cognitive impairment in depression: a systematic review and meta-analysis. *Psychol Med*. 2014;44(10):2029-2040. doi:10.1017/S0033291713002535
50. Brissos S, Dias VV, Kapczinski F. Cognitive performance and quality of life in bipolar disorder. *Can J Psychiatry*. 2008;53(8):517-524. doi:10.1177/070674370805300806
51. Ceylan D, Akdede BB, Bora E, et al. Neurocognitive functioning during symptomatic states and remission in bipolar disorder and schizophrenia: A comparative study. *Psychiatry Res*. 2020;292. doi:10.1016/j.psychres.2020.113292
52. Dybedal GS, Tanum L, Sundet K, Gaarden TL, Bjølseth TM. Cognitive side-effects of electroconvulsive therapy in elderly depressed patients. *Clin Neuropsychol*. 2014;28(7):1071-1090. doi:10.1080/13854046.2014.958536
53. Obbels J, Verwijk E, Vansteelandt K, et al. Long-term neurocognitive functioning after electroconvulsive therapy in patients with late-life depression. *Acta Psychiatr Scand*. 2018;138(3):223-231. doi:10.1111/acps.12942
54. Nuninga JO, Claessens TFI, Somers M, et al. Immediate and long-term effects of bilateral electroconvulsive therapy on cognitive functioning in patients with a depressive disorder. *J Affect Disord*. 2018;238:659-665. doi:10.1016/j.jad.2018.06.040

55. Borsboom D, Deserno MK, Rhemtulla M, et al. Network analysis of multivariate data in psychological science. *Nat Rev Methods Prim*. 2021;1(1). doi:10.1038/s43586-021-00055-w
56. Cassano GB, Rucci P, Frank E, et al. The mood spectrum in unipolar and bipolar disorder: arguments for a unitary approach. *Am J Psychiatry*. 2004;161(7):1264-1269. doi:10.1176/APPI.AJP.161.7.1264
57. Biondi M, Picardi A, Pasquini M, Gaetano P, Pancheri P. Dimensional psychopathology of depression: detection of an "activation" dimension in unipolar depressed outpatients. *J Affect Disord*. 2005;84(2-3):133-139. doi:10.1016/S0165-0327(02)00103-9
58. Angst J. The bipolar spectrum. *Br J Psychiatry*. 2007;190(MAR.):189-191. doi:10.1192/BJP.BP.106.030957
59. Vieta E, Phillips ML. Deconstructing bipolar disorder: a critical review of its diagnostic validity and a proposal for DSM-V and ICD-11. *Schizophr Bull*. 2007;33(4):886-892. doi:10.1093/SCHBUL/SBM057
60. Salagre E, Vieta E. Precision psychiatry: Complex problems require complex solutions. *Eur Neuropsychopharmacol*. 2021;52:94-95. doi:10.1016/J.EURONEURO.2021.07.003
61. Wittenborn AK, Rahmandad H, Rick J, Hosseinichimeh N. Depression as a systemic syndrome: mapping the feedback loops of major depressive disorder. *Psychol Med*. 2016;46(3):551. doi:10.1017/S0033291715002044
62. Kapur S, Phillips A, Insel T. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry*. 2012;17(12):1174-1179. doi:10.1038/MP.2012.105
63. van Eeden WA, van Hemert AM, Carlier IVE, et al. Basal and LPS-stimulated inflammatory markers and the course of individual symptoms of depression. *Transl Psychiatry*. 2020;10(1). doi:10.1038/s41398-020-00920-4
64. Fried EI, Nesse RM, Zivin K, Guille C, Sen S. Depression is more than the sum score of its parts: individual DSM symptoms have different risk factors. *Psychol Med*. 2014;44(10):2067-2076. doi:10.1017/S0033291713002900
65. Fried EI, Nesse RM. Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. *BMC Med*. 2015;13(1):72. doi:10.1186/s12916-015-0325-4
66. Bringmann LF, Vissers N, Wichers M, et al. A Network Approach to Psychopathology: New Insights into Clinical Longitudinal Data. *PLoS One*. 2013;8(4):e60188. doi:10.1371/JOURNAL.PONE.0060188
67. Pedrelli P, Blais MA, Alpert JE, Shelton RC, Walker RSW, Fava M. Reliability and validity of the Symptoms of Depression Questionnaire (SDQ). *CNS Spectr*. 2014;19(6):535-546. doi:10.1017/S1092852914000406
68. Fried EI. The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *J Affect Disord*. 2017;208(October 2016):191-197. doi:10.1016/j.jad.2016.10.019
69. Fried EI, Robinaugh DJ. Systems all the way down: Embracing complexity in mental health research. *BMC Med*. 2020;18(1):4-7. doi:10.1186/s12916-020-01668-w

70. Wichers M, Schreuder MJ, Goekoop R, Groen RN. Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology. *Psychol Med*. 2019;49(3):380-387. doi:10.1017/S0033291718002064
71. Witherington DC. The dynamic systems approach as metatheory for developmental psychology. *Hum Dev*. 2007;50(2-3):127-153. doi:10.1159/000100943
72. Jeronimus BF. Dynamic System Perspectives on Anxiety and Depression. http://www.rug.nl/research/portal. Accessed April 27, 2022.
73. Wichers M, Wigman JTW, Bringmann LF, de Jonge P. Mental disorders as networks: some cautionary reflections on a promising approach. *Soc Psychiatry Psychiatr Epidemiol*. 2017;52(2):143-145. doi:10.1007/s00127-016-1335-z
74. Dejonckheere E, Mestdagh M, Houben M, et al. Complex affect dynamics add limited information to the prediction of psychological well-being. *Nat Hum Behav*. 2019;3(5):478-491. doi:10.1038/s41562-019-0555-0
75. Semkovska M, McLoughlin DM. Objective cognitive performance associated with electroconvulsive therapy for depression: A systematic review and meta-analysis. *Biol Psychiatry*. 2010;68(6):568-577. doi:10.1016/j.biopsych.2010.06.009
76. Bos EH, Ten Have M, van Dorsselaer S, Jeronimus BF, de Graaf R, de Jonge P. Functioning before and after a major depressive episode: pre-existing vulnerability or scar? A prospective three-wave population-based study. *Psychol Med*. 2018;48(13):2264-2272. doi:10.1017/S0033291717003798
77. Frasch K, Bullacher C, Jäger M, et al. Effects of symptom reduction and psychotropic medication on cognitive impairment in depression. *Psychopathology*. 2009;42(1):59-66. doi:10.1159/000187635
78. McClintock SM, Choi J, Deng Z De, Appelbaum LG, Krystal AD, Lisanby SH. Multifactorial determinants of the neurocognitive effects of electroconvulsive therapy. *J ECT*. 2014;30(2):165-176. doi:10.1097/YCT.0000000000000137

Summary

Mood disorders are complex disorders with multiple factors determining their onset and underlying etiopathogenetic mechanisms. It is a great challenge to incorporate this complex, multi-layered thinking in research methods. In this doctoral thesis, we present the results of six original studies and one meta-analysis of the literature that were all aimed at gaining an increased insight in the etiopathogenesis and/or course of MDD and BD. The three studies from Part 1 (Chapter 2 to 4) were founded on more conventional and widely used research paradigms, in which mood disorders are seen as discretely delineated categories and mean sum-scores of (mood or inflammatory) outcome variables are compared on the group-level. Part 2 (Chapter 5 and 6) took upon a different approach by studying dynamics of individual symptoms both at individual-level, as well as aggregated outcomes, to gain a more fine-grained insight in the symptom dynamics of MDE. Part 3 (Chapter 7 and 8) shows another application of a combined individual-level and group-level approach, this time to study the cognitive outcomes after treatment with ECT in MDE.

This doctoral thesis provides only a glimpse of the complexity of mood disorders. Specifically, we pointed towards the highly complex and idiosyncratic dynamic behavior of symptoms over time and confirmed hereby that the use of sum-scores has important limitations. Furthermore, we showed the importance of individual-level analyses (in addition to group-level analyses) to enlarge the clinical translatability of research findings. Not only depressive symptoms should be considered as highly dynamically complex entities of a larger pathology, but also context, bio-psycho-social, and temperamental variables could be included in further analyses which will likely make a worthwhile endeavor.

Samenvatting

Stemmingsstoornissen zijn complexe aandoeningen waarvan het ontstaan en het verloop wordt bepaald door verschillende, onderling interagerende, factoren. De studies beschreven in dit proefschrift zijn allen gericht op het verkrijgen van meer inzicht in het verloop en de etiopathogenese van stemmingsstoornissen. Het proefschrift is opgebouwd uit drie delen. In het eerste deel wordt de symptomatologie en het verloop van een depressie over de tijd onderzocht binnen een klinische (residentiële) behandelsetting op basis van routine outcome monitoring data. Daarna wordt de rol van het tryptofaan metabolisme in de etiopathogenese van bipolaire stemmingsstoornissen (en de bijhorende cognitieve symptomen) onderzocht. Deel 2 gaat dieper in op het verloop van depressie over de tijd maar onderzoekt dit op symptoom niveau. Het verloop van individuele symptomen over de tijd en de dynamische samenhang met andere symptomen wordt onderzocht, zowel op het niveau van de individuele patiënt als op groepsniveau. In deel 3 gaan we dieper in op de cognitieve symptomen bij patiënten met een stemmingsstoornis, meer specifiek onderzoeken we de cognitieve (neven)effecten na een behandeling met electroconvulsietherapie. Waar voorgaand ECT onderzoek dit voornamelijk onderzocht op groepsniveau, combineren we in deel 3 individuele analyses met groepsanalyses om zo een genuanceerde blik op cognitieve (neven)effecten van ECT te bieden.

Dit proefschrift verhult slechts een klein aspect van de complexiteit in onderzoek naar stemmingstoornissen aan. Meer specifiek komt het idiosyncratische karakter van depressie symptomen over de tijd naar voren, wat vragen doet stellen bij het wijdverspreide gebruik van som scores in onderzoeksanalyses. Verder tonen we de potentiële meerwaarde van analyses op individueel niveau (naast groepsniveau analyses). Niet enkel depressie symptomen gedragen zich idiosyncratisch over de tijd, ook andere bio-psycho-sociale variabelen van depressie vertonen een dynamisch verloop en onderlinge interacties die vermoedelijk patiënt-specifiek zijn. Het onderzoeken van stemmingsstoornissen als dynamische entiteiten, bestaande uit verschillende interagerende factoren vormt een veelbelovend pad voor verder onderzoek.

Abbreviations

3-HK	3-hydroxykynurenine
AMI	Autobiographic Memory Interview
BACS-C	Brief Assessment of Cognition in Schizophrenia-Symbol Coding
BBP	Bruto binnenlands product
BD	Bipolar disorder
BDI	Beck Depression Inventory
CPT-IP	Continuous Performance Task-Identical Pairs
CSF	Cerebrospinal fluid
D-KEFS	Delis-Kaplan Executive Function System
DSM	Diagnostic and Statistical Manual of Mental Disorders
DTW	Dynamic time warp
ECT	Electroconvulsive therapy
ICD	International Classification of Diseases and Related Health Problems
ISBD-BANC	International Society for Bipolar Disorders Battery for assessment of Neurocognition
GDP	Gross domestic product
HRSD	Hamilton Rating Scale for Depression
HVLT-R	Hopkins Verbal Learning Test – Revised
KA or KYNA	Kynurenic acid
KYN	Kynurenine
LCGA	Latent class growth analysis
LMM	Linear mixed models
LNS	Letter Number Span
MANSA	Manchester Short Assessment of Quality of Life
MDD	Major depressive disorder
MDE	Major depressive episode
MINI	Mini-International Neuropsychiatric Interview
MoCA	Montreal Cognitive Assessment
NMDA	N-methyl-D-aspartate
QA	Quinolinic acid
PROTECT	Prediction of treatment response to electroconvulsive therapy study
RCI	Reliable Change Index
ROM	Routine Outcome Monitoring
RDoC	Research Domain Criteria
SDST	Symbol Digit Substitution Test
SQ-48	Symptom Questionnaire-48
TRP	Tryptophan
TRYCAT	Tryptophan catabolites
VAR	Vector autoregressive model
WHO	World Health Organization

Dankwoord

Dit proefschrift draag ik op aan de deelnemers van de verschillende studies vermeld in dit proefschrift maar ook aan die van de TMS studie waarin ik actief betrokken was. Jullie engageerden zich voor deelname aan klinisch onderzoek in een moeilijke periode in jullie leven. Met dit werk hoop ik een bijdrage geleverd te hebben aan de zoektocht naar het verloop en oorzaken van stemmingsstoornissen.

Prof dr. Sabbe, het was een hele eer om onderzoek te mogen doen onder uw vleugels. Onderzoek uitvoeren met een brede blik op de wereld en nooit de verbinding met de kliniek verliezen, het zal iets zijn dat ik in mijn verdere carrière hoog in het vaandel zal dragen. Dank voor de kansen die u me heeft gegeven en voor de vele boeiende discussies samen.

Prof. dr. Giltay, Erik, samenwerken met jou was een puur plezier. Je eindeloze enthousiasme voor wetenschappelijk onderzoek werkt aanstekelijk. Ik denk met een warm gevoel terug aan de stageperiode in LUMC Leiden, waar we naast onderzoeks- ook klinische collega's werden. Veel dank om Lize en mij te introduceren in jullie stad, Gouda. Ralph en jij hebben nog een gastronomisch bezoekje aan Brussel tegoed!

Prof. dr. Morrens, Manuel, zoals een goede werkbegeleider dat betaamt, was je er steeds op cruciale momenten. Bedankt om steeds mee na te denken over oplossingen en te relativeren waar dat nodig was.

Een speciaal woord van dank aan de Universiteit Antwerpen en de directie van het UPC Duffel, en in het bijzonder aan Roeland Depreitere. Dank voor de mogelijkheden om dit onderzoek tot een goed einde te brengen. Veel dank ook aan de klinische behandelteams van UPC Duffel en de ROM testpsychologen (o.a. Ellen Clé en Sylvie) die betrokken waren bij de onderzoeksprojecten in dit proefschrift. Merci aan de onderzoeksstagiairs en jobstudenten (Laura, Nathalie en Richard) voor hun hulp bij het verwerken van de ruwe data.

Dank aan mijn collega PhD studenten met wie werk en plezier hand in hand gingen. Oli, het werd snel duidelijk dat onze basispijlers van een goed PhD traject gelijkgestemd waren: goede koffie en buitenlandse congressen. Praag, Warschau, New York: een systematische kwaliteitsanalyse van de beschikbare rooftops werd grondig uitgevoerd en bracht onze vriendschap naar een (toepasselijk toch?) hoger niveau. Vaak gieren van het lachen, maar daarnaast ook hard werken. Niet zelden verloren we ons in een lichtelijk megalomane gedachtegang (de ene al wat meer dan de andere) over hoe we de psychiatrie van de toekomst zouden vormen. Ik kijk uit naar onze volgende projecten samen. Linda en Seline:

dank om me verder te laten bouwen op jullie onderzoeksprojecten. Van jullie leerde ik het meest. Ik mis onze brainstorm momenten over onderzoeksvragen en statistische analyses en hoop dat we in onze drukke levens hier nog een gaatje voor zullen blijven vinden. Mirella, door jouw veerkracht en doorzettingsvermogen kon het Herculesproject rekenen op een stevige basis. Ik ben blij dat ik dit werk samen met jou verder kon uitbouwen. Eva, dank voor het meedenken en nalezen van de laatste versie van het proefschrift. Jobbe, Kawtar, Jean-Baptiste, Mariana, Katrien en Annelies, het was fijn samenwerken met jullie!

Dank aan het voltallige Sinaps team (onder leiding van prof. dr. Manuel Morrens, prof. dr. Violette Coppens, en dr. Livia De Picker) voor de samenwerking. In het bijzonder veel dank aan Alysia, Tina, Sarah, Anneke en Elien voor de actieve bijdrage aan de patiënteninclusies! Merci ook aan Katrien (Steurs) voor de goede coördinatie (en de leuke babbels tussendoor). Merci lieve Ingrid, moeder Sinaps. Altijd paraat en daarnaast een topchef in kerstperioden voor het voltallige onderzoeksteam!

Veel dank aan de (nog niet vermelde) co-auteurs van de artikels. Roos van der Mast en Tom Birkenhäger: jullie snelle en inzichtgevende suggesties hebben me steeds stof tot nadenken gegeven. Esmée en Jasmien: jullie inzichten waren erg waardevol en betekenden een grote meerwaarde voor de uiteindelijke wetenschappelijke 'output'. Ook veel dank aan Didier Schrijvers, Alexander van Nuijs, Ella Roelant en Marijke Dejaeger voor het meedenken. Eiko, it was such an honour and pleasure to work with you. Thank you for your sharing your valuable insights into the field of symptomics and dynamic research approaches.

Mijn collega's van de Vlaamse Vereniging voor Assistenten Psychiatrie. De gemeenschappelijke strijd voor een kwaliteitsvolle opleiding en de liefde voor onze (toekomstige) job bracht ons samen. Ik ben er zeker van dat onze wegen nog vaak zullen kruisen.

Bedankt aan mijn collega-psychiaters van het UPC KULeuven voor hun interesse en steun. Een groot woord van dank aan dr Sabien Wyckaert voor de samenwerking en de patiëntenverwijzingen voor de TMS studie. Ik kijk uit naar onze uitbouw van de multi-disciplinaire raadpleging voor patiënten met een bipolaire stoornis op campus Kortenberg. Team MIC: merci voor de fijne samenwerking van het afgelopen jaar en om mij de laatste weken hier en daar wat te ontzien. Dank aan iedereen van de ECT onderzoeksgroep AcCENT voor de steun in deze laatste periode voor de verdediging. Ik kijk uit naar onze toekomstige samenwerking.

Merci aan de collega's van UZ Brussel (in het bijzonder Dieter Zeeuws en Chris Baeken) en UPC KULeuven (Chris Bervoets en Choi Deblieck) voor de samenwerking voor de TMS studie.

Bedankt aan de interne juryleden, prof. dr. Geert Dom, prof. dr. Roy Remmen en prof. dr. Paul Van Royen om de voortgang van dit doctoraat mee op te volgen en te ondersteunen. Bedankt aan prof. dr. Stephan Claes en prof. dr. Robert Schoevers voor jullie interesse en deelname als extern jurylid.

Bedankt aan de getalenteerde 'fille du block' Silke (a.k.a. Sire) voor het ontwerpen van de coverafbeelding. Verder ook bedankt aan Hanneke Meulenbroek voor de taalsuggesties.

Merci aan de leukste (Brusselse en Tongerse) vriendinnen die steeds voor de nodige ontspanning klaarstonden.

Dank aan mijn (schoon)zussen en schoonbroers. Aan de leukste neefjes en nichtjes: al ravottend met jullie vergeet ik helemaal wat er nog 'to do' is.

Merci aan mijn schoonouders, Peter & Martine, om de schilderwerken in ons appartement over te nemen wanneer bij ons het verzadigingspunt bereikt was.

Papa, als geen ander liet je ons zien hoe waardevol het is om je steeds te laten verwonderen en plezier te vinden in je werk. Ik ben blij dat ik nu eens een boek aan jou kan geven voor in je boekenkast (ipv andersom). Mama, onze werkijver en doorzettingsvermogen die hebben we van jou. Bedankt voor je steun (in werkelijk alles wat we doen) en voor je luisterend oor.

En als laatste, mijn favoriete compagnon de route, Lize. Onze gezamenlijke avonturen zijn de leukste. Wat een jaar hebben we achter de rug! Beiden een nieuwe job, verhuizen & verbouwen, een trouwfeest, PhD. In 2023 graag een vitesse lager 😊 Je t'aime!

KAAT HEBBRECHT

SCHAARBEEK, 1030, BELGIUM 0473238951

DETAILS

Smaragdlaan
Schaarbeek, 1030
Belgium
0473238951
kaat.hebbrecht@hotmail.com

Date of birth
27/04/1989

Nationality
Belg

LANGUAGES

Nederlands

Engels

Frans

EXPERIENCE

Adult Psychiatrist at UPC KULeuven
November 2021 — Present

- Medium Intensive Care Unit
- Bipolar disorder outpatient clinic
- Electroconvulsive therapy

Psychiatric resident at UPC KULeuven
August 2014 — June 2021

- LUMC (Leiden, NL): consultation-liaison psychiatry and outpatient clinic (2020-2021)
- UPC KULeuven (Kortenberg, BE): psychotic disorder ward, bipolar disorder mood ward (2016-2017)
- UPC KULeuven (Gasthuisberg, BE): consultation-liaison psychiatry, crisisward, outpatient clinic for mood and anxiety (2014-2016)

PhD researcher at University of Antwerp
April 2017 — November 2020

- Routine Outcome Monitoring in Depression and the study of Personalized Symptom Dynamics in Depression. Financiering: Herculesstichting (FWO)
- Continuous Theta Burst Stimulation as an add-on Treatment for Bipolar Depression (ClinicalTrials.gov Identifier: NCT03603561): multi-center study (UPC Duffel, UZ Brussel, UPC KULeuven)

EDUCATION

Master in Medicine, KULeuven , Leuven
September 2010 — June 2014

Magna Cum Laude

Bachelor in Medicine, UHasselt, Hasselt
September 2007 — June 2010

Cum Laude

Secondary School , VIIO Humaniora , Tongeren
September 2000 — September 2007

COURSES

Workshop for Early Career Scientists, ECNP
March 2021

Postgraduate in Psychodynamic Psychotherapy, KULeuven
2016 — 2020

Statistical courses, Antwerp Doctoral School
2017 — 2020

Course Scientific Writing, European Psychiatric Association (EPA)
2019

Summerschool on Affective Neuroscience, BCNBP
2015

VOLUNTARY WORK

Establishment of a medical center, Nalou, Mali
2010

TEACHING AND SUPERVISION SCIENTIFIC WORK

Supervision of master theses (Master in Medicine), Universiteit Antwerpen
2019 — July 2021

Teaching communication skills (Master in Medicine), Universiteit Antwerpen
November 2017 — January 2018

EXTRA-CURRICULAR ACTIVITIES

Member at Vlaamse Vereniging voor Psychiatrie
2018 — Present

President at Vlaamse Vereniging Assistenten Psychiatrie
2018 — 2020

Editor at De Jonge Psychiater
2018 — June 2021
www.dejongepsychiater.nl

ORAL PRESENTATIONS

Hebbrecht K. (2022). ECT en cognitieve neveneffecten. Oral presentation presented at the sectie-vergadering ECT (VVP), Kortenberg, Belgium.

Hebbrecht K. (2022). Een genuanceerde blik op cognitieve (neven)effecten van ECT. Oral presentation presented at the Vlaams Geestelijk Gezondheidszorgcongres, Antwerpen, Belgium.

Hebbrecht K. (2022). Psychiatrie in Nederland en België: verschillen in de opleiding tot psychiater . Oral presentation presented at the conference of the Orde van den Prince, Nijmegen, The Netherlands.

Hebbrecht K. (2022). Motivationele gespreksvoering en psychiatrische casuïstiek in de apothekerpraktijk. Oral presentation presented at the LOK Zuid-Limburg (apothekers), Hasselt, Belgium.

Hebbrecht K (2021). ROM and the search for clinical and biological endophenotypes in depression: tackling heterogeneity. Research seminar present at the Jelgersmalezing, Leiden, The Netherlands.

Hebbrecht K. (2021). Routine Outcome Monitoring en klinisch gestuurd onderzoek in de Vlaamse psychiatrie. Oral presentation presented at the Vlaams Geestelijk Gezondheidszorgcongres, Antwerpen, Belgium,.

Hebbrecht K (2021). The role of kynurenine metabolites in cognitive dysfunction in bipolar disorder. Poster presented at the ECNP workshop for early career scientists, online.

Hebbrecht K. (2020). Dynamic time warp analysis of Hamilton Depression Rating Scale (HDRS) items in depressed inpatients. Oral presentation presented at the NVVP Voorjaarscongres, online.

Hebbrecht K. (2019). cTBS in bipolar depression. Oral presentation presented at the Brainstimulation conference, Gent, Belgium.

Hebbrecht K. (2019). Symptom profile and clinical course of inpatients with unipolar versus bipolar depression. Poster presented at the EPA congress, Warschau, Poland.

★ PUBLICATIONS

Lambrichts S, Vansteelandt K, Hebbrecht K, Wagenmakers MJ, Oudega ML, Obbels J, van Exel E, Dols A, Bouckaert F, Schrijvers D, Verwijk E, Sienaert P. Which residual symptoms predict relapse after successful electroconvulsive therapy for late-life depression? J Psychiatr Res. 2022 Oct;154:111-116

Hebbrecht K, Morrens M, Giltay EJ, van Nuijs ALN, Sabbe B, van den Ameele S. The Role of Kynurenines in Cognitive Dysfunction in Bipolar Disorder. Neuropsychobiology. 2022;81(3):184-191

Hebbrecht K, Skorobogatov K, Giltay EJ, Coppens V, De Picker L, Morrens M. Tryptophan Catabolites in Bipolar Disorder: A Meta-Analysis. Front Immunol. 2021 May 19;12:667179

Hebbrecht, K., Stuivenga, M., Birkenhäger, T., Morrens, M., Fried, E.I., Sabbe, B., & Giltay, E.J. (2020) Understanding personalized dynamics to inform precision medicine: a dynamic time warp analysis of 255 depressed inpatients. *BMC medicine.* 18(1), 400.

Hebbrecht K, Giltay EJ, Birkenhäger TK, Sabbe B, Verwijk E, Obbels J, Roelant E, Schrijvers D, Van Diermen L. (2020) Cognitive change after electroconvulsive therapy in mood disorders measured with the Montreal Cognitive Assessment. *Acta Psychiatr Scand.* 142(5):413-422.

Hebbrecht K, Stuivenga M, Birkenhäger T, van der Mast RC, Sabbe B, Giltay EJ. (2020) Symptom Profile and Clinical Course of Inpatients with Unipolar versus Bipolar Depression. *Neuropsychobiology*. 79(4-5):313-323.

Hebbrecht K, Pattyn T. (2020) Psychiatrie van de toekomst: sterk verhoogd risico op ongunstige gebeurtenissen na een psychiatrische opname. *Tijdschrift voor Psychiatrie* 61. 8, 588 - 588

Cools O., & Hebbrecht K. (2019) De opleiding psychiatrie in Vlaanderen door de ogen van de arts-specialist in opleiding. *Tijdschrift voor psychiatrie.* 3, 153–158.

Hebbrecht, K., Morrens, M., Neels, H., Roosens, L., & Sabbe, B. G. C. (2018) Pharmacokinetic evaluation of the aripiprazole (once-monthly) injection for the treatment of bipolar disorder. *Expert opinion on drug metabolism & toxicology.* 10, 999–991005.

van Diermen, L., Hebbrecht, K., Schrijvers, D., Sabbe, B. C. G., Fransen, E., & Birkenhäger, T. K. (2018) The Maudsley Staging Method as predictor of electroconvulsive therapy effectiveness in depression. *Acta psychiatrica Scandinavica.* 6, 605–614.

Cools O, Hebbrecht K, Coppens V, Roosens L, De Witte A, Morrens M, Neels H, Sabbe B (2018) Pharmacotherapy and nutritional supplements for seasonal affective disorders: a systematic review. *Expert Opin Pharmacother*. 19(11):1221-1233.