INSTRUMENTS TO OPTIMIZE MENTAL HEALTH(CARE)

FRÉDÉRIQUE VAN KRUGTEN

Instruments to Optimize Mental Health(Care)

Frédérique van Krugten

ISBN: 978-94-6361-654-6
Layout and printing: Optima Grafische Communicatie, Rotterdam, The Netherlands

Instruments to Optimize Mental Health(Care)

Instrumenten ter optimalisatie van de geestelijke gezondheid(szorg)

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. A.L. Bredenoord

en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op
donderdag 31 maart 2022 om 10:30 uur

door

Frédérique Cornélie Wilhelmina van Krugten
geboren te Delft

Erasmus University Rotterdam

Promotiecommissie

Promotor:	Prof.dr. W.B.F. Brouwer
Overige leden:	Prof.dr. J.M. Cramm Prof.mr.dr. S.M.A.A. Evers Prof.dr. H.G.J.M. Vermetten
Copromotor:	Dr. L. Hakkaart-van Roijen

Contents

1 General introduction

Background

Mental health problems are common [1]. The World Health Organization (WHO) estimated that one in every four people will be affected by one or more mental health problems at some time during their life [2]. Mental health problems, such as depression and anxiety, are generally characterized by a combination of abnormal thoughts, emotions, behaviors, and difficulties with interpersonal relationships [3] and are associated with a wide range of adverse consequences that can take place on different levels. On an individual level, mental health problems significantly impact people's health, functioning, and quality of life [1, 4, 5], and, especially for more severe conditions, may lead to a lower life expectancy [6]. In addition, the people close to the person suffering from a mental health problem (e.g., partners, family members, friends) can be significantly affected. This is not only since witnessing a loved one in need or experiencing potential relational issues has been shown to directly affect the well-being of a significant other (sometimes labeled the family effect), but also since significant others often offer first emotional and practical support to people with mental health problems (sometimes labeled the caregiving effect) [7]. Especially in more severe circumstances, providing informal support and care to people with mental health problems, often on a regular and intensive basis, can have a profound effect on these caregivers and is, amongst others, associated with burnout and a decreased quality of life [8, 9]. In addition to the suffering of the people experiencing mental health problems and the people close to them, the societal burden of mental health problems is profound. Mental health problems are, for instance, associated with substantial healthcare costs and productivity losses [10, 11]. In particular circumstances and mental conditions, the educational and criminal justice systems may also be affected, as well as a general sense of safety in society [12-16]. Hence, the burden of disease related to mental health problems is large, both on an individual and societal level.

In light of the high prevalence and associated burden of mental health problems, national health systems strive to deliver accessible and high-quality mental healthcare. However, given the limited resources available, it is vital that the care provided is necessary, effective and cost-effective, i.e., that scarce resources are used to obtain most (health) benefits. Ensuring this is challenging and relates to both the organization of the (mental) healthcare system and the type of treatments offered within the system. One of the persistent challenges in the organization of the mental healthcare system is the on average lengthy patient pathway through the layers of mental healthcare [17]. For some patients, especially the ones with complex and severe mental health problems, it can take a long time before they receive the level and type of care they require, often implying that they received ineffective or at least insufficiently effective

care until that moment [18-21]. One of the ways to enhance the efficiency of mental healthcare delivery and reduce suffering, therefore, is to optimize the pathways with which patients "flow" through the mental healthcare system in order to ensure that they receive appropriate care in a timely manner. Optimization of these patient pathways involves, among other things, the early identification of patients in need of highly specialized treatment to facilitate a timely referral to appropriate treatment settings. The underlying assumption here is that a quicker provision of appropriate care to these patients enhances health outcomes and cost-effectiveness of care or may even reduce costs. However, instruments that identify patients with mental health problems in need of highly specialized care are largely lacking and their development is, therefore, warranted.

In order to be able to demonstrate that patients in need of highly specialized mental healthcare benefit from quicker referral to appropriate care levels, but also to assess the effectiveness and cost-effectiveness of mental health interventions in general, it is pivotal that outcomes of treatments can be adequately evaluated. Mental healthcare interventions typically aim to reduce symptoms and complaints and improve patients' quality of life. In order to measure quality of life, also in the context of assessing cost-effectiveness, generic health-related quality of life instruments have been developed that are used to evaluate a wide variety of interventions. However, concerns have been raised regarding their suitability for use in the context of (parts of) mental healthcare [22]. More specifically, it has been suggested that generic health-related quality instruments are, in certain conditions and levels of severity, not sufficiently sensitive to relevant (changes in) dimensions of quality of life relevant in the context of mental health. To assess the efficiency with which services are provided, also to inform the optimization of patient pathways in mental healthcare, it is vital that appropriate instruments are in place. This thesis addresses the issue on the lengthy average patient pathway through the layers of mental healthcare and the issue on quality of life measurement in the mental health field with the overarching question: *Can we develop instruments to optimize mental health(care)?*

Before specifying the research questions of this thesis, first the organization of mental healthcare and the models of service delivery as well as how outcomes of mental healthcare interventions can be evaluated will be introduced to provide the context in which this research was performed.

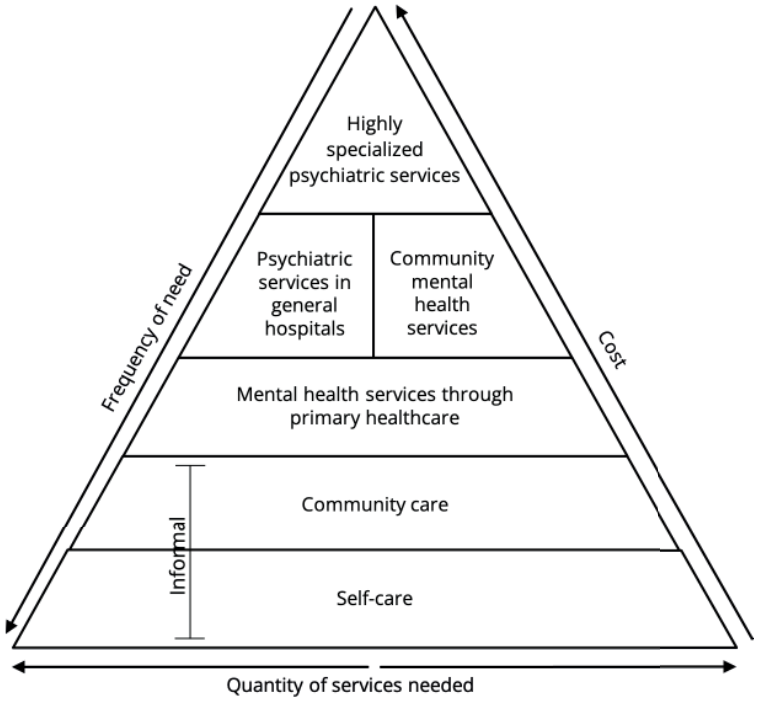
The organization of mental healthcare

The provision of effective mental health interventions is vital in reducing the burden of mental health problems [2]. Mental health interventions, such as antidepressant treatment and psychotherapy, are provided through mental healthcare services. The way in which these services are organized is likely to have a profound impact on the effectiveness and efficiency of the mental health interventions offered [23].

Although the optimal way of organizing mental health services is largely subject to a country's epidemiological, cultural, political, and economic climate, the WHO provided recommendations on the key elements for the successful provision of mental health services [23]. These recommendations were summarized and visually represented in a framework as illustrated in Figure 1.1. As indicated by the framework, most mental health problems can be managed through self-care and informal community mental health services (i.e., care provided by local community members). When the mental health needs require additional support beyond that which can be offered by informal care, intervention should be provided through more formalized, professional services. These formalized services include, generally speaking, in descending order of frequency of need and ascending order of costs and specialization: primary, secondary, and tertiary mental healthcare services. Within these echelons of formalized mental health services, the majority of people are ideally treated in primary care services. Secondary mental healthcare is the specialized support provided to patients with mental health problems that cannot be adequately managed by primary care services [23]. As illustrated by the "Optimal Mix of Services Pyramid" in Figure 1.1, secondary mental healthcare is ideally provided by community mental health services and psychiatric services in general hospitals. Tertiary mental healthcare, also known as highly specialized mental healthcare, is the mental healthcare offered by highly trained mental health specialists with competence in a certain field to individuals with mental health disorders that require intervention beyond that provided by secondary mental healthcare [24, 25]. Highly specialized mental healthcare treatment is often provided through academic or top-clinical healthcare institutes due to the needed degree of personnel experience, management, security, and resources [25].

In the context of rising demands and scarce resources for mental healthcare [2, 26], it is pivotal to optimize the organization of service delivery in order to ensure that available services are used to their best effect. This, among other things, involves the optimization of patient pathways through the different layers (i.e., echelons) of the mental healthcare system in order to enhance the provision of the right treatment to the right patient at the right time. Currently, patients who ultimately end up in tertiary

Figure 1.1. World Health Organization's "Optimal Mix of Services Pyramid". Adapted from the World Health Organization's (WHO) "Optimal Mix of Services Pyramid" [23].



care facilities often have had a long history of, arguably inadequate, care consumption in lower echelons [18-21]. Avoiding such suboptimal patient pathways could reduce suffering, increase the clinical effectiveness and cost-effectiveness of mental healthcare and may even prove to be cost-saving.

The stepped and matched care models of service delivery attempt to optimize service delivery and refer to the way in which patients flow through the network of formalized mental health services. Within the stepped care model of service delivery, patients generally receive the least intensive and costly treatment, such as guided self-help or psychoeducational classes, as a first-line treatment [27]. More intensive treatments, provided in secondary and tertiary mental healthcare settings, are reserved for patients who benefit insufficiently from initial treatment [27]. The stepped care model is designed to manage large numbers of patients with relatively mild mental health problems at lower costs and simplify patient care pathways through standardization of procedures. In contrast, in the matched care model of service delivery, pre-treatment patient characteristics are used to match patients to the treatment that is likely to be most beneficial to them [28]. Hence, within this model, patients who are expected

not to benefit from the least intensive treatment can be referred to more intensive treatments without initially receiving the least intensive treatment available. The matched care model, therefore, allows the needs of patients with severe and complex conditions to be met in a timelier manner by allocating them directly to more intensive treatments. In addition, since delay in establishing the optimal treatment intensity has been associated with partial recovery and chronicity [29, 30], the matched care model has the potential of being more (cost-) effective than the stepped care model in patients who are predictively in need of intensive treatment. However, a challenge related to the matched care model is the identification of accurate individual patient indicators with which patients in need of highly specialized mental healthcare can be adequately identified. Information on such patient indicators collected as part of the routine assessment procedure could facilitate an early, prioritized referral to more intensive treatments when needed. Hence, one of the key elements for the successful application of the matched care model of service delivery is the early identification of patients in need of highly specialized mental healthcare. This raises the question of whether it is possible to identify pre-treatment patient characteristics that could facilitate the systematic and standardized early identification of patients with a highly specialized mental healthcare need. This question is addressed in the first part of this thesis (Chapter 2-6).

Evaluating the outcomes of mental healthcare

In order to be able to demonstrate that patients in need of highly specialized mental healthcare indeed benefit from quicker referrals and more tailored treatment pathways, but also to assess the effectiveness and cost-effectiveness of mental health interventions in general, outcomes of treatments need to be evaluated adequately. To facilitate this, the routine use of standardized outcome instruments to systematically measure the progress and outcomes of the provided care is becoming an important part of daily clinical practice around the world [31]. Despite the interest in, and sometimes even obligation to, monitor and evaluate outcomes of the provided care, questions about what to measure and how to measure appropriate outcomes remain subject to critical debate [22].

One area of debate in the context of outcome measurement in the mental health field concerns the measurement of health-related quality of life. Health-related quality of life is a widely accepted outcome measure, also in the mental health field [32]. Although an agreed upon definition is lacking, health-related quality of life is generally viewed as a subjective and multidimensional construct that signifies the impact of a

condition and its treatment on the quality of life of an individual [33]. In addition to evaluating the clinical effectiveness, health-related quality of life is also used in the context of economic evaluations of mental healthcare interventions [22]. Economic evaluations compare the costs and benefits of a healthcare intervention relative to one or more alternatives in order to assess its value for money [34]. Different types of economic evaluations exist, which can be distinguished in the way they quantify the health benefits of the compared interventions. A popular and often recommended type of economic evaluation is cost-utility analysis (CUA), in which these benefits are expressed as quality-adjusted life-years (QALYs). The QALY combines length of life (i.e., the number of remaining years a person is expected to live) and the health-related quality of life experienced during those years.

The most commonly used method for measuring health-related quality of life and therefore to inform the 'Q' in the QALY is through the administration of standardized self-completion instruments (i.e., questionnaires) [35]. On such instruments, respondents are asked to indicate how they score on a number of dimensions (i.e., domains) of quality of life, such as physical health, mental well-being, social relations, and daily activities. Based on the provided responses, often, a score can be calculated that represents the level of quality of life of an individual. For QALY calculation, this score, often called a utility score, normally reflects the relative preference for a specific health state. Such utility values are anchored at 1 (full health) and 0 (death) and are typically pre-defined and obtained in a representative sample of the general population through choice-based techniques such as the time trade-off method and discrete choice experiments [34].

Given the relevance and necessity of measuring quality of life, numerous standardized quality of life instruments have been developed and these instruments are used frequently, also in the mental health field. An important distinction between these instruments is whether they are generic, domain-specific or disease-specific. Generic instruments focus on quality of life dimensions that are relevant to all conditions, whereas domain-specific and disease-specific instruments are sensitive to dimensions of quality of life that are relevant to a specific group of conditions (e.g., mental health problems in general) or one particular condition (e.g., major depressive disorder). Despite the multitude of available quality of life instruments, concerns have been raised regarding the suitability of existing quality of life instruments for use in the context of (parts of) mental healthcare [36, 37]. While for many disease and domain-specific outcome measures utility scores are lacking and comparability (even within the mental health domain) is questioned, frequently used generic quality of life instruments have been shown to be insufficiently sensitive to certain relevant (changes in) dimensions

of quality of life relevant in the context of mental health [37, 38]. It has been suggested that this may be the result of the focus on physical health in the development of generic quality of life instruments, which consequently may have constrained the coverage of the dimensions perceived important to the quality of life of people with mental health problems [39]. In other words, the inability of frequently used generic quality of life instruments to adequately capture the benefits of mental healthcare interventions might be related to the content validity of these instruments. Previous work by Connell et al. [40, 41] identified seven dimensions that are important in defining quality of life of people with mental health problems, which raises the question whether existing quality of life instruments sufficiently cover these seven dimensions. Insufficient coverage of these dimensions may lead to inadequate estimations of the benefit of mental health interventions, and may ultimately lead to suboptimal resource allocations. Such concerns also raise the question whether it is necessary and possible to develop new quality of life measures that capture all relevant quality of life dimensions in the context of mental health. This question is addressed in the second part of this thesis (Chapter 7-9).

Research questions

This introduction started with the notion that, in light of rising demands and scarce resources available for mental healthcare, there is a need to enhance the efficiency of mental healthcare in order to improve outcomes and safeguard their sustainability. This thesis addresses this need by exploring whether it is possible to develop instruments to improve the (evaluation of the) efficiency of mental healthcare. More specifically, the research presented in this thesis focuses on (1) the development and validation of instruments to optimize pathways of patients in need of highly specialized mental healthcare and (2) the development, validation and valuation of a mental health-related quality of life instrument. Below, the research questions this thesis addresses are listed in more detail. The answers to these research questions, as provided in this thesis, aim to contribute to improving the (evaluation of the) efficiency of mental healthcare and ultimately to reducing the burden of mental health problems.

The research questions addressed in this thesis are:

1. Which individual patient indicators could facilitate the systematic and standardized early identification of patients in need of highly specialized mental healthcare?
2. How to develop and what are the psychometric properties of diagnosis-specific instruments that facilitate the systematic and standardized early identification of patients in need of highly specialized mental healthcare?

3. How to develop and what are the psychometric properties of an instrument that facilitates the systematic, early and standardized, transdiagnostic identification of patients in need of highly specialized mental healthcare?
4. Which quality of life instruments are currently used to measure the outcomes of mental healthcare interventions and what is their content validity?
5. Is it possible to develop a psychometrically sound mental health-related quality of life instrument?
6. Can we derive a valuation set with utility scores for a new mental health-related quality of life instrument?

Outline of this thesis

Chapters 2 to 6 report on the development and psychometric evaluation of instruments to optimize the efficiency of patient pathways of patients with a highly specialized mental healthcare need. **Chapter 2** reports the findings of a systematic literature review that was carried out to identify pre-treatment patient characteristics that could facilitate the systematic and standardized identification of patients with major depressive disorder in need of highly specialized care. **Chapter 3** describes a concept mapping study that was performed to complement the indicators derived from the systematic literature review (Chapter 2) with clinical expertise. Based on the results from Chapters 1 and 2, an instrument aimed at aiding clinicians in the early identification of patients with major depressive disorder in need of highly specialized care was developed, the Decision Tool Unipolar Depression (DTUD). The development and psychometric evaluation of the DTUD are reported in **Chapter 4**. Building on the theoretical foundations of, and insights from, the development of the DTUD, **Chapter 5** describes the development and psychometric evaluation of the Decision Tool Anxiety Disorders, OCD and PTSD (DTAOP). The development and psychometric evaluation of the diagnosis-specific Decision Tools suggested that the allocation of patients to highly specialized mental healthcare settings, in general, may be guided by a core set of transdiagnostic patient factors. **Chapter 6** therefore reports on the development and psychometric evaluation of a transdiagnostic decision tool that facilitates the systematic and standardized early identification of patients with a need for highly specialized mental healthcare.

Chapters 7 to 9 report on the development, psychometric evaluation and valuation of a mental health-related quality of life instrument, the Mental Health Quality of Life questionnaire (MHQoL). **Chapter 7** assesses to what extent currently used quality of life instruments cover the dimensions perceived important to the quality of life of

people with mental health problems. **Chapter 8** describes the development and first psychometric evaluation of the MHQoL. **Chapter 9** reports on the estimation of a value set with standard scores to generate health state utility values for QALY calculations, making the MHQoL suitable for use in cost-utility evaluations.

Finally, in **Chapter 10** the main findings of this dissertation are summarized, discussed and interpreted in the context of research and policy. In addition, recommendations for further research and policy are provided. Please note that Chapters 2-9 are based on papers published in or submitted to international peer-reviewed journals and can therefore be read independently.

2

Indicators of patients with major depressive disorder in need of highly specialized care: A systematic review

van Krugten FCW
Kaddouri M
Goorden M
van Balkom AJLM
Bockting CLH
Peeters FPML
Hakkaart-van Roijen L

PLOS ONE. 2017;12(2):e0171659.

Abstract

Objectives

Early identification of patients with major depressive disorder (MDD) that cannot be managed by secondary mental health services and require highly specialized mental healthcare could enhance need-based patient stratification. This, in turn, may reduce the number of treatment steps needed to achieve and sustain an adequate treatment response. The development of a valid tool to identify patients with MDD in need of highly specialized care is hampered by the lack of a comprehensive understanding of indicators that distinguish patients with and without a need for highly specialized MDD care. The aim of this study, therefore, was to systematically review studies on indicators of patients with MDD likely in need of highly specialized care.

Methods

A structured literature search was performed on the PubMed and PsycINFO databases following PRISMA guidelines. Two reviewers independently assessed study eligibility and determined the quality of the identified studies. Three reviewers independently executed data extraction by using a pre-piloted, standardized extraction form. The resulting indicators were grouped by topical similarity, creating a concise summary of the findings.

Results

The systematic search of all databases yielded a total of 7,360 references, of which sixteen were eligible for inclusion. The sixteen papers yielded a total of 48 unique indicators. Overall, a more pronounced depression severity, a younger age of onset, a history of prior poor treatment response, psychiatric comorbidity, somatic comorbidity, childhood trauma, psychosocial impairment, older age, and a socioeconomically disadvantaged status were found to be associated with proxies of need for highly specialized MDD care.

Conclusions

Several indicators are associated with the need for highly specialized MDD care. These indicators provide easily measurable factors that may serve as a starting point for the development of a valid tool to identify patients with MDD in need of highly specialized care.

Introduction

Major depressive disorder (MDD) is one of the most prevalent psychiatric disorders [42, 43] and is projected to be the leading cause of disease burden in high-income countries by 2030 [44]. MDD presents many treatment challenges, not the least of which is the subset of patients with depression that is refractory to secondary mental health services. Often, these patients receive inadequate, too low-intensity treatment in secondary mental health services [18-21], which is associated with a longer treatment course [30], an increased risk of suicide [45-47] and substantial societal costs [29, 48].

Early identification of patients with MDD who cannot be managed by secondary services and require highly specialized care could enhance need-based patient stratification. This, in turn, may reduce the number of treatment steps needed to achieve and sustain an adequate treatment response, and may subsequently benefit the quality of life of patients. To date, validated tools to facilitate need-based patient stratification are rarely used in psychiatric practice. This is in marked contrast to other areas of medicine such as oncology [49-52], in which patient stratification on the basis of clinical presentation plays an important role in treatment planning from the time of diagnosis.

The development of a validated tool to identify patients with MDD in need of highly specialized care during the diagnostic phase after referral is hampered by the lack of a comprehensive understanding of the indicators that distinguish patients with and without a need for highly specialized MDD care. There are several reviews available which summarize the studies on factors associated with a recurrent or persistent clinical course [53-55] for which more intensive treatment is indicated [56]. However, to date none have focused on the factors associated with a broad range of unfavorable clinical outcomes, thereby preventing the construction of an overall picture of the indicators of patients with MDD in need of highly specialized care. Therefore, the aim of this study is to systematically review studies on indicators of patients with MDD likely in need of highly specialized care.

Methods

Definition of terms

For the purpose of this study, primary mental healthcare is defined as the care provided to people with mental health problems within the primary care setting. Second-

ary mental healthcare is delivered primarily through community mental health services and psychiatric services in general hospitals, and refers to the more specialized support provided to patients with mental health needs that cannot be supported by primary care services [23]. Highly specialized mental healthcare, also commonly referred to as tertiary mental healthcare, is defined as specialized intervention delivered by highly-trained staff with specific expertise in a given field to individuals with mental health problems that cannot be treated with sufficient result by either primary or secondary mental health services [24, 25]. Finally, the term "indicators" is used to refer to clinical characteristics and risk factors that may aid clinicians in the identification of the subgroup of patients with MDD likely in need of highly specialized care.

Expert input and proxy indicators of need for highly specialized MDD care

Prior to performing the structured literature search, the Decision Tool Unipolar Depression Consortium was formed comprising thirteen leading MDD experts from six independent psychiatric specialized and highly specialized mental healthcare clinics across the Netherlands. The consortium of experts assisted with refining the research question and provided guidance for the conduct of the literature search. In the absence of studies directly examining clinical and socio-demographic factors associated with a need for highly specialized MDD care, proxy indicators had to be identified. In a digital survey, consortium members and a number of other qualified domain experts were asked to define terms by which (the clinical course of) patients with MDD in need of highly specialized care can be described (hereafter named proxy indicators of need for highly specialized MDD care). All domain experts were required to have specialist expertise regarding the research question, as evidenced by the fact that they were either active as a clinician or researcher in the field of depression. Ultimately, via existing national depression networks, 134 experts were approached, 67 of whom participated in the study. After an analysis of the concepts submitted by the experts, four high-frequency proxy indicators of need for highly specialized MDD care were selected. The proxy indicators of need for highly specialized MDD care that were selected for the purpose of this review included: "Treatment-Resistant", "Chronic", "Recurrent", and "Persistence of Severity".

Eligibility criteria

Studies were selected for review if they met the following inclusion criteria:

1. Published in English or Dutch, related to humans and full-text available;
2. Published between January 2000 and January 2015;
3. The study design was either a randomized controlled trial, case-control study, cross-sectional study or cohort study;

4. The study was an investigation of (a group of) adult psychiatric patients (aged 18 and over) with MDD as their primary diagnosis according to the Diagnostic and Statistical Manual of Mental Disorders (DSM)-III [57], DSM-III-R [58], DSM-IV [59], DSM-IV-TR [60], DSM-5 [61], International Classification of Diseases (ICD)-9 [62], ICD-10 [63] or Research Diagnostic Criteria (RDC) [64];
5. The main outcome variable used was one of the four following proxy indicators of need for highly specialized MDD care: "Treatment-Resistant", "Chronic", "Recurrent", and "Persistence of Severity";
6. One of the aims of the study was to identify clinical and/or socio-demographic factors that discriminate MDD patients with a proxy of need for highly specialized care (i.e. "cases") from those without a proxy indicator of need for highly specialized care (i.e. "non-cases").

This study is restricted to indicators of patients with MDD in need of highly specialized care, which can be assessed during the diagnostic phase after referral. Hence, no papers that solely reported on physiological, neurobiological, or genetic factors were eligible for inclusion. Furthermore, since the aim was to identify indicators of patients with an unfavourable treatment course treated in secondary mental health services who may benefit from highly specialized care, we excluded studies focusing exclusively on participants from primary care populations or the general population.

Data sources and search strategy

To identify studies reporting indicators of patients with MDD in need of highly specialized care, a structured literature search was performed on the PubMed (National Library of Medicine) and PsycINFO (Ovid) databases following PRISMA guidelines [65]. The search for published primary articles was conducted on January 15, 2015 and was restricted to articles written in English or Dutch, published between January 1, 2000 and January 15, 2015, related to humans and for which the full text was available. Search terms were chosen based on the proxy indicators of need for highly specialized MDD care as defined by domain experts. The Medical Subject Headings (MeSH) of relevance to this review included the following search terms: "depressive disorder", "depression" and "depressive disorder, treatment-resistant". In addition, keywords were searched within the title or full-text. Keywords included: "chronic", "chronic depression", "chronicity", "recurrent", "recurrent depression", "recurring", "severe", "severe depression" and "severity". A complete list of search strategies can be found in Appendix 2.1. We did not register a systematic review protocol.

Study selection

Prior to examining all articles identified through the primary search, two reviewers independently screened a random sample of 66 titles and abstracts whilst blinded to authors and journal titles, and reached strong agreement (Cohen's $\kappa = 0.85$) using an Excel workbook designed for this purpose [66]. They then independently screened all records whilst still blinded to authors and journal titles. Full papers were retrieved for all references that had been judged as potentially eligible and were examined independently by two researchers. Disagreements were resolved by discussion or through third party adjudication.

Data abstraction

Three reviewers independently executed data extraction by using an Excel-based, pre-piloted, standardized extraction form. Disagreements were resolved by discussion between the reviewers. The following characteristics of the studies were coded: (1) general study characteristics (author, year of publication, country); (2) characteristics of the study population (sample size, age of inclusion, mean age, number of MDD patients with and without a proxy of need for highly specialized care); (3) design of the study (case-control, cross-sectional or longitudinal); (4) depression measure and proxy of need for highly specialized MDD care (e.g. treatment-resistant, recurrence); (5) clinical and/or socio-demographic factors on which MDD patients with and without a proxy of need for highly specialized care significantly differed. If results from a multivariable regression analysis were available, then those findings were included rather than bivariate results. If results from several regression models were presented, only results from the model with the largest number of predictors were used. The purpose of this review was to identify, rather than quantify, the factors associated with proxies of need for highly specialized care. Thus, a meta-analysis was not performed and data were synthesized in a narrative review. In a consensus-building process, the experts categorized the abstracted indicators by topical similarity, creating a concise summary of the findings. The resulting categories were identified as the overarching indicators of patients with MDD in need of highly specialized care, provided that the direction of association between the indicators grouped within the category and proxies of need for highly specialized care was consistent. Within the categories grouping indicators with opposite directions (e.g. low and high educational level), subcategories of indicators with a consistent direction of association were identified as the indicators of patients with MDD in need of highly specialized care.

Quality assessment

Two reviewers independently evaluated the methodological quality of the included studies using the 14-item National Heart, Blood and Lung Institute (NHBLI) Quality

Assessment Tool for Observational Cohort and Cross-Sectional Studies [67] or the 12-item NHBLI Quality Assessment Tool for Case-Control Studies [68]. Each of the items was scored as "yes", "no", "not reported" or "not applicable" on the basis of the information provided in the paper. Disagreements were resolved by discussion or through third party adjudication. A quality score, expressed as a percentage of the maximum possible score, was calculated for each study.

Results

Study selection

The systematic search of all databases yielded a total of 7,360 references. Duplicates were checked and excluded (n=1,388). Title and abstract screening resulted in the exclusion of a further 5,917 papers. Main reasons for exclusion were that papers: had a design other than a randomized controlled trial, case-control study, cross-sectional study or cohort study; had an initial population or control group other than subjects with MDD as their primary diagnosis; had an aim other than the identification of clinical and/or socio-demographic factors that discriminate patients on the basis of a proxy of need for highly specialized care. Full texts of the remaining 55 papers were obtained for detailed review. Thirty-nine papers were excluded following full text screening. Sixteen papers fulfilled the eligibility criteria and were incorporated into the review. Details of the study selection process are provided in Figure 2.1.

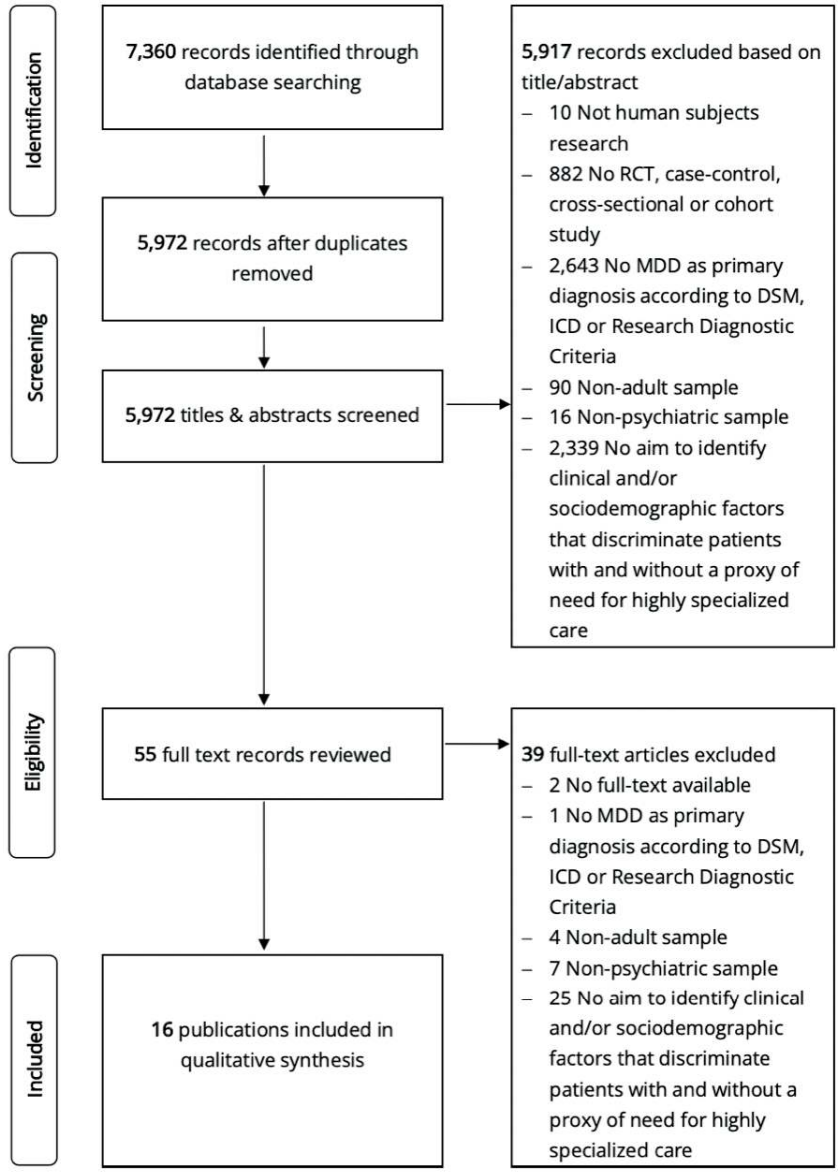
Study characteristics

The general characteristics of the included papers are presented in Table 2.1. Six papers focused on treatment-resistant depression [69-74], four on chronic depression [75-78], five on recurrence in depression [79-83], and one on persistence of severity [84]. The vast majority of included papers utilized cross-sectional data. Most of the included studies were conducted in the United States (n=5) and Europe (n=8) with the remainder in Asia (n=3).

Methodological quality of the included studies

The overall quality scores are presented in Table 2.1; quality scores for the separate NHBLI criteria are presented in Appendix 2.2. The overall quality scores ranged from 36% [72-74] to 86% [79]. The main issues with included papers were the lack of sample size justification and the lack of repeated exposure assessment. The research question and study population were clearly defined in the majority of the included studies.

Figure 2.1. Flow chart of study selection process.



RCT, Randomized controlled trial; MDD, major depressive disorder; DSM, Diagnostic and Statistical Manual of Mental Disorders; ICD, International Classification of Diseases.

Indicators of need of highly specialized care

Table 2.2 lists the indicators of patients with a depression in need of highly specialized care. The sixteen papers yielded a total of 48 unique clinical and socio-demographic factors on which MDD patients with and without a proxy of need for highly specialized

Table 2.1. General characteristics of the included studies.

Outcome examined and study	Population	Sample size (C/NC)	Mean age C/NC (SD)	Diagnostic criteria	Study design (follow-up length)	Country	NHBLI Quality Score (%)
Treatment-resistance							
Kaplan et al. 2000 [69]	Outpatients from a university clinic	40 (20/20)	47 (*not given*) / 45 (*not given*)	DSM-IV and ICD-10	Nested case-control	USA	50
Souery et al. 2007 [70]	Outpatients and inpatients from specialist referral centers	702 (356/346)	50.5 (14.1) / 51.5 (14.6)	DSM-IV	Cross-sectional	European countries	57
Amital et al. 2008 [71]	Outpatients from community psychiatric clinics	107 (42/65)	54.7 (16.3) / 49.6 (16.2)	ICD-10	Cross-sectional	Israel	43
Dudek et al. 2010 [72]	Outpatients from psychiatric clinics	1,051 (570/481)	47 (11) / 46 (11)	DSM-IV-TR	Cross-sectional	Poland	36
Takahashi et al. 2013 [73]	Outpatients from university clinics	62 (35/27)	38.74 (9.42) / 39.07 (9.19)	DSM-IV	Cross-sectional	Japan	36
Takahashi et al. 2013 [74]	Outpatients from university clinics	66 (35/31)	35.94 (8.93) / 38.00 (8.42)	DSM-IV	Cross-sectional	Japan	36
Chronicity							
Riso et al. 2003 [75]	Outpatients from a university mood disorders unit	69 (42/27)	39.3 (10.3) / 39.1 (10.3)	DSM-IV	Cross-sectional	USA	43
Gilmer et al. 2005 [76]	Outpatients from primary or psychiatric care sites	1,380 (293/1,087)	41.9 (13.5) / 39.7 (13.0)	DSM-IV	Cross-sectional	USA	43
Wiersma et al. 2009 [77]	Subjects from the community, primary care settings, and specialized mental healthcare facilities	1,204 (395/809)	42.4 (11.8) / 39.7 (12.3)	DSM-IV	Cross-sectional	NL	50
Wiersma et al. 2011 [78]	Subjects from the community, primary care settings, and specialized mental healthcare facilities	1,002 (312/690)	40.5 (12.2) / 43.2 (11.8)	DSM-IV	Cross-sectional	NL	50

Table 2.1. General characteristics of the included studies. (continued)

Outcome examined and study	Population	Sample size (C/NC)	Mean age C/NC (SD)	Diagnostic criteria	Study design (follow-up length)	Country	NHBLI Quality Score (%)
Recurrence							
Melartin et al. 2004 [79]	Secondary-level care psychiatric outpatients and inpatients	198 (76/122)	41.0 (11.1)[a]	DSM-IV	Longitudinal (18 months)	Finland	86
Solomon et al. 2004 [80]	Outpatients and inpatients from academic medical centers	290 (143/147)	39 (15)[a]	RDC	Longitudinal (15 years)	USA	71
Bos et al. 2005 [81]	Female outpatients	50 (30/20)	*Not given*	DSM-IV	Cross-sectional	NL	43
Hollon et al. 2006 [82]	Outpatients from primary and psychiatric care sites	1,426 (1,061/365)	41.2 (13.2) / 38.9 (13.4	DSM-IV	Cross-sectional	USA	50
Gerrits et al. 2014 [83]	Subjects from the community, primary care settings, and specialized mental healthcare facilities	*Not given* (292/ *not given*)	43.4 (12.8)[b]	DSM-IV	Longitudinal (4 years)	NL	71
Persistence of severity							
Lamers et al. 2011 [84]	Subjects from the community, primary care settings, and specialized mental healthcare facilities	789 (19%/81%)	41.8 (12.0)[a]	DSM-IV	Longitudinal (1 year)	NL	64

C/NC = Cases/Non-Cases; SD = Standard Deviation; NHBLI = National Heart, Blood and Lung Institute; DSM = Diagnostic and Statistical Manual of Mental Disorders; ICD = International Classification of Diseases; RDC = Research Diagnostic Criteria; USA = United States of America; NL = The Netherlands.

[a] For the overall sample.

[b] For the overall sample (n=1,122), including anxiety patients.

care significantly differ. In general, the abstracted clinical and socio-demographic factors could be grouped into the following seven categories: depression severity, onset and (treatment) course, comorbid psychopathology, somatic comorbidity, childhood trauma, psychosocial functioning, and socio-demographics. Each of the abstracted indicators will be discussed in the following sections.

Table 2.2. Indicators of patients with a depression in need of highly specialized care.

Indicator
Depression severity
Greater (baseline [84]) depressive symptom severity [77, 79]
Current suicidal risk [70]
Higher rates of melancholic features [70]
Higher levels of rumination [78]
Onset and (treatment) course
Younger age of onset [72, 82]
Longer time since first onset [82]
History of prior suicide attempts [76]
Shorter current episode [82]
Less likely to meet criteria for chronic depression [82]
More than three previous depressive episodes [72]
Fewer prior episodes of depression [76]
Lack of remission or partial remission after the previous depressive episode [72]
Nonresponse to first antidepressant treatment lifetime [70]
Comorbid psychopathology
A higher number of comorbid psychiatric disorders [79]
Comorbid (generalized [76]) anxiety disorder [69, 70, 77]
Higher levels of chronic PTSD [69]
More symptoms of bipolarity [72]
Higher scores on the MMPI-2 subscales [69]
Lower levels of extraversion on the NEO-FFI [73, 78]
Lower levels of reward dependence on the TCI-125 [74]
Lower levels of self-directedness on the TCI-125 [74]
Higher levels of harm avoidance on the TCI-125 [74]
Higher levels of impaired autonomy on the YSQ [75]
Higher levels of disconnection and rejection on the YSQ [75]
Higher levels of overvigilance on the YSQ [75]
Higher levels of external locus of control on the SMS [78]
Higher levels of neuroticism on the NEO-FFI [84]
Somatic comorbidity
Greater general medical comorbidity [76]
Worse physical health function [76]

Table 2.2. Indicators of patients with a depression in need of highly specialized care. (continued)

Indicator
Lower physical quality of life [76]
Severe neck, chest and abdominal pain [83]
A higher number of pain locations [83]
Higher severity of pain [83]
Childhood trauma
Higher prevalence of childhood trauma [77]
Greater levels of childhood emotional abuse [69]
Higher levels of trauma sequelae [69]
Psychosocial functioning
Worse work function and social adjustment [76]
Impaired psychosocial functioning [80]
Lower quality of life [76]
Socio-demographics
Older age [76, 81, 82]
Less education [76]
Higher educational level [81]
Lower monthly household income [76]
No private insurance [76]
Unemployment [76]
Prior job loss [71]
A greater likelihood of being Black as opposed to white or other [76]
A greater likelihood of being Hispanic as opposed to non-Hispanic [76]

MMPI-2 = Minnesota Multiphasic Personality Inventory-2; NEO-FFI = NEO Five-Factor Inventory; TCI-125 = Cloninger's 125-question Temperament and Character Inventory; YSQ = Young Schema Questionnaire; SMS = Self-Mastery Scale.

Depression severity

A more pronounced depression severity, whether operationalized by the number or type of symptoms, has consistently been shown to be associated with proxies of need for highly specialized MDD care. Three studies reported a correlation between higher levels of (baseline [84]) depressive symptomatology and proxies of need for highly specialized MDD care [77, 79]. In addition, the presence of certain symptoms such as current suicidal risk [70], an increased likelihood of melancholic features [70], and higher levels of rumination [78] were found to be associated with proxies of need for highly specialized MDD care.

Onset and (treatment) course

Patients with a proxy of need for highly specialized care were found to have an earlier age of onset of the first major depressive episode [72, 82], and subsequently reported

a longer time since first onset of MDD [82]. In addition, factors denoting a history of poor treatment response such as nonresponse to first antidepressant received [70], and lack of remission or partial remission after the previous depressive episode [72] were found to be associated with proxies of need for highly specialized care. Inconsistent results were found for the number of prior episodes of depression [72, 76].

Comorbid psychopathology

There have been several studies that examined the association between comorbid psychopathology and proxies of need for highly specialized care. Melartin et al. [79] found that the presence of a higher number of comorbid psychiatric disorders in general increases the risk of recurrence. In addition, the following specific comorbid psychiatric disorders were found to be associated with proxies of need for highly specialized care: chronic PTSD [69], (generalized [76]) anxiety disorder [69, 70, 77], and more symptoms of bipolarity [72]. Furthermore, relations between psychopathological dimensional personality traits and proxies of need for highly specialized care have been found repeatedly. In two closely related articles, Takahashi et al. [73, 74] reported that high scores for harm avoidance, low scores for reward dependence, low scores for self-directedness, and low scores for extraversion are personality dimensions in patients with treatment-resistant depression. In addition, Kaplan and Klinetob [69] reported that patients with treatment-resistant depression had clinically significant elevations on the Minnesota Multiphasic Personality Inventory-2 (MMPI-2 [85]) subscales hypochondriasis, depression, hysteria, psychopathic deviate, paranoia, psychasthenia, schizophrenia and social introversion. Further, higher levels of impaired autonomy [75], higher levels of disconnection and rejection [75], higher levels of overvigilance [75], higher levels of external locus of control [78], and higher levels of neuroticism [84] have been linked to proxy indicators of need for highly specialized care.

Somatic comorbidity

Increased general medical comorbidity [76], severe neck, chest and abdominal pain [83], a higher number of pain locations [83] and higher severity of pain [83] were found to be associated with proxies of need for highly specialized care. Subsequently, lower levels of physical health function [76] and a lower physical quality of life [76] have been linked to proxy indicators of need for highly specialized care.

Childhood trauma

Two studies [69, 77] examined the relationship between childhood trauma and a proxy indicator of need for highly specialized care. In a sample of 1,230 individuals, Wiersma et al. [77] examined the relationship between retrospective reports of childhood life events and childhood trauma and the risk of chronicity of MDD in adulthood.

They found that a reported history of multiple childhood traumas, such as emotional neglect, psychological abuse, physical abuse, and sexual abuse, was associated with chronicity of depression. Kaplan and Klinetob [69] similarly found that patients with treatment-resistant depression reported more emotional abuse and experienced current-day trauma sequelae when compared to treatment responders.

Psychosocial functioning

Two of the included studies [76, 80] reported that patients with a proxy of need for highly specialized care were more likely to exhibit impaired functioning in areas such as work, relationships and leisure. Moreover, a poorer quality of life, as operationalized by the Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q [86]), was found to be associated with a proxy of need for highly specialized care [76].

Socio-demographics

Many studies examined the associations between socio-demographic factors and proxies of need for highly specialized care. Three papers [76, 81, 82] reported an association between older age and a proxy of need for highly specialized care. In addition, individuals with a proxy of need for highly specialized care were found to be socioeconomically disadvantaged when compared to individuals without a proxy of need for highly specialized care [76]. Contrasting findings were found for educational level. One study [76] reported that patients with a lower level of education exhibited greater chronicity than patients with a higher level of education. By contrast, the study by Bos et al. [81] found the reverse: patients with a history of recurrent depression were more highly educated compared to individuals with a single episode.

Discussion

The aim of this systematic review was to identify indicators of patients with MDD in need of highly specialized care. Overall, a more pronounced depression severity, a younger age of onset, a history of prior poor treatment response, psychiatric comorbidity, somatic comorbidity, childhood trauma, psychosocial impairment, older age, and a socioeconomically disadvantaged status were found to be associated with proxies of need for highly specialized MDD care.

To our knowledge, this is the first systematic literature search that comprehensively covers the factors associated with a broad range of unfavorable clinical outcomes in patients with MDD for which more intensive treatment is indicated [57]. To date, reviews solely summarized factors associated with one of the proxy indicators of need

for highly specialized care [53-55], thereby preventing the construction of an overall picture of the indicators of patients with MDD in need of highly specialized care. Our systematic and comprehensive review allows the delineation of this subgroup of patients, makes them identifiable, and thus adds to the process of further professionalizing and improving quality in the mental healthcare sector.

This study has several limitations. First, this study does not shed light on the efficacy of highly specialized care in meeting patients' treatment needs. Although highly specialized care has been demonstrated to improve clinical outcomes in patients with complex and severe conditions in other areas of medicine [87], the net benefit of highly specialized care in patients with MDD has not yet been studied. However, the evaluation of the impact of highly specialized care on patient outcomes in this population is of utmost importance and should therefore be addressed in future studies. Second, the focus of this systematic review was to identify indicators that could be easily assessed in routine clinical practice, specifically during the diagnostic phase after referral. This resulted in the exclusion of papers solely reporting on physiological, neurobiological, and genetic patient factors, making it possible that other indicators with strong evidence for a need for highly specialized MDD care have been missed. Third, due to considerable heterogeneity of populations, sample sizes, range of predictors, outcomes and statistical analyses no quantitative synthesis of the results in a meta-analysis could be performed. Fourth, since the aim of this study was to assess the current state of research on indicators of patients with an unfavourable treatment course treated in secondary mental health services who may benefit from highly specialized care, we also included studies with a heterogeneous mixture of patients from the community and from primary and psychiatric care sites, as they contained a subgroup of psychiatric patients. This may have influenced the results, as recent studies suggest that the determinants and nature of the long-term course of depression of subjects from the community and primary sites differentiates from that of patients from psychiatric care sites [88-90]. However, an additional qualitative synthesis of the results in the subset of studies exclusively reporting on patients with MDD treated in psychiatric care sites did not alter the results, suggesting that the associations between indicators of need for highly specialized care are similar for psychiatric and non-psychiatric patients.

On the basis of this review, we posit the primary importance of the following nine indicators of patients with MDD in need of highly specialized care: a more pronounced depression severity, a younger age of onset, a history of prior poor treatment response, psychiatric comorbidity, somatic comorbidity, childhood trauma, psychosocial impairment, older age, and a socioeconomically disadvantaged status. It should be noted,

however, that these indicators alone are not likely to justify referral to highly specialized mental healthcare programs. Rather, in combination with one another they may provide healthcare practitioners with a guideline for determining the need for highly specialized care. Future research should explore how the identified set of indicators can facilitate the early identification of patients with MDD in need of highly specialized care. In addition, we believe that advances in the development of a valid tool to identify patients with MDD in need of highly specialized care during the diagnostic phase after referral will need to be based on more refined, better operationalized indicators. Furthermore, while the identified indicators have received the strongest support in the literature, this may partly be due to the fact that they have received more research attention. It is therefore possible that other characteristics of patients in need of highly specialized care may theoretically be very important, but have not yet been sufficiently researched. Hence, in accordance with evidence-based medicine [91], this set of characteristics should be critically appraised, refined, and, if necessary, complemented by clinical expertise before applying review findings to clinical practice. The identified set of indicators may therefore serve as a starting point for the development of a valid tool to identify patients with MDD in need of highly specialized care during the diagnostic phase after referral. This may ultimately facilitate early detection and assist clinicians in selecting the most appropriate treatment option in a given clinical situation, thereby reducing the functional impact and socioeconomic burden of MDD.

Appendix 2.1 – Search strategies

Table 2.1.1. Search strategy Pubmed (NLM).

#	Searches
1	("Depressive Disorder"[Mesh major topic] OR "Depression"[Mesh major topic]) AND ("Depressive Disorder, Treatment-Resistant"[Mesh major topic] OR chronic[ti] OR "chronic depression" OR chronicity[ti] OR complex[ti] OR "complex depression" OR complexity[ti] OR complicated[ti] OR recurrent[ti] OR "recurrent depression" OR recurring[ti] OR severe[ti] OR "severe depression" OR severity[ti])
2	limit 1 to full text
3	limit 2 to human
4	limit 3 to english or Dutch language
5	limit 101 to yr="2000 -Current"

Table 2.1.2. Search strategy PsycINFO (Ovid).

#	Searches
1	"depressive disorder".ti,ab.
2	"disorder, depressive".ti,ab.
3	"disorders, depressive".ti,ab.
4	"neurosis, depressive".ti,ab.
5	"depressive neuroses".ti,ab.
6	"depressive neurosis".ti,ab.
7	"neuroses, depressive".ti,ab.
8	"depression, endogenous".ti,ab.
9	"endogenous depression".ti,ab.
10	"endogenous depressions".ti,ab.
11	"depressive syndrome".ti,ab.
12	"depressive syndromes".ti,ab.
13	"syndrome, depressive".ti,ab.
14	"syndromes, depressive".ti,ab.
15	"depression, neurotic".ti,ab.
16	"depressions, neurotic".ti,ab.
17	"neurotic depression".ti,ab.
18	"neurotic depressions".ti,ab.
19	"melancholia".ti,ab.
20	"melancholias".ti,ab.
21	"unipolar depression".ti,ab.
22	"depression, unipolar".ti,ab.
23	"depressions, unipolar".ti,ab.
24	"unipolar depressions".ti,ab.
25	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24
26	depression.ti,ab.

Table 2.1.2. Search strategy PsycINFO (Ovid). (continued)

#	Searches
27	depressions.ti,ab.
28	"depressive symptoms".ti,ab.
29	"depressive symptom".ti,ab.
30	"symptoms, depressive".ti,ab.
31	"emotional depression".ti,ab.
32	"depression, emotional".ti,ab.
33	"emotional depressions".ti,ab.
34	26 or 27 or 28 or 29 or 30 or 31 or 32 or 33
35	"depressive disorder, treatment-resistant".ti,ab.
36	"depressive disorders, treatment resistant".ti,ab.
37	"treatment-resistant depressive disorders".ti,ab.
38	"treatment-resistant depressive disorder".ti,ab.
39	"therapy-resistant depression".ti,ab.
40	"therapy resistant depression".ti,ab.
41	"therapy-resistant depressions".ti,ab.
42	"treatment resistant depression".ti,ab.
43	"depression, treatment resistant".ti,ab.
44	"resistant depression, treatment".ti,ab.
45	"treatment resistant depressions".ti,ab.
46	"refractory depression".ti,ab.
47	"depression, refractory".ti,ab.
48	chronic.ti.
49	"chronic depression".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
50	chronicity.ti.
51	recurrent.ti.
52	"recurrent depression".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
53	recurring.ti.
54	severe.ti.
55	"severe depression".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
56	severity.ti.
57	25 or 34
58	35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57
59	57 and 58
60	limit 98 to full text
61	limit 99 to human
62	limit 100 to english or dutch language
63	limit 101 to yr="2000 -Current"

Appendix 2.2 – Quality assessment

Table 2.2.1. Quality assessment for observational cohort and cross-cectional studies.

Study	Quality assessment criterion[a]														Quality score (%)
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Souery et al. 2007	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	No	Yes	CD	NA	Yes	57
Amital et al. 2008	Yes	Yes	Yes	Yes	No	No	No	Yes	No	No	Yes	CD	NA	No	43
Dudek et al. 2010	Yes	No	NR	Yes	No	No	No	No	Yes	No	Yes	CD	NA	Yes	36
Takahashi et al. 2013	Yes	No	NR	Yes	No	No	No	Yes	Yes	No	Yes	CD	NA	No	36
Takahashi et al. 2013	Yes	No	NR	Yes	No	No	No	Yes	Yes	No	Yes	CD	NA	No	36
Riso et al. 2003	Yes	No	NR	Yes	No	No	No	Yes	Yes	No	Yes	CD	NA	Yes	43
Gilmer et al. 2005	Yes	Yes	NR	Yes	No	No	No	Yes	Yes	No	Yes	CD	NA	No	43
Wiersma et al. 2009	Yes	Yes	NR	Yes	No	No	No	Yes	Yes	No	Yes	CD	NA	Yes	50
Wiersma et al. 2011	Yes	Yes	NR	Yes	No	No	No	Yes	Yes	No	Yes	CD	NA	Yes	50
Melartin et al. 2004	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	CD	Yes	Yes	86
Solomon et al. 2004	Yes	Yes	NR	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	NR	71
Bos et al. 2005	Yes	No	Yes	CD	No	No	No	Yes	Yes	No	Yes	CD	NA	Yes	43
Hollon et al. 2006	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	No	No	No	NA	Yes	50
Gerrits et al. 2014	Yes	Yes	NR	Yes	No	Yes	Yes	Yes	Yes	No	Yes	CD	Yes	Yes	71
Lamers et al. 2011	Yes	No	NR	Yes	No	Yes	Yes	Yes	Yes	No	Yes	CD	Yes	Yes	64

CD = Cannot Determine; NA = Not Applicable; NR = Not Reported.

[a] (1) Was the research question or objective in this paper clearly stated? (2) Was the study population clearly specified and defined? (3) Was the participation rate of eligible persons at least 50%? (4) Were all the subjects selected or recruited from the same or similar populations? Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants? (5) Was a sample size justification, power description, or variance and effect estimates provided? (6) For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured? (7) Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed? (8) For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)? (9) Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants? (10) Was the exposure(s) assessed more than once over time? (11) Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants? (12) Were the outcome assessors blinded to the exposure status of participants? (13) Was loss to follow-up after baseline 20% or less? (14) Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?

Table 2.2.2. Quality assessment for case–control studies.

Study	Quality assessment criterion[a]												Quality score (%)
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	
Kaplan et al. 2000	Yes	No	No	Yes	Yes	No	CD	CD	Yes	Yes	Yes	No	50

CD = Cannot Determine.

[a] (1) Was the research question or objective in this paper clearly stated and appropriate? (2) Was the study population clearly specified and defined? (3) Did the authors include a sample size justification? (4) Were controls selected or recruited from the same or similar population that gave rise to the cases (including the same time-frame)? (5) Were the definitions, inclusion and exclusion criteria, algorithms or processes used to identify or select cases and controls valid, reliable, and implemented consistently across all study participants? (6) Were the cases clearly defined and differentiated from controls? (7) If less than 100 percent of eligible cases and/or controls were selected for the study, were the cases and/or controls randomly selected from those eligible? (8) Was there use of concurrent controls? (9) Were the investigators able to confirm that the exposure/risk occurred prior to the development of the condition or event that defined a participant as a case? (10) Were the measures of exposure/risk clearly defined, valid, reliable, and implemented consistently (including the same time period) across all study participants? (11) Were the assessors of exposure/risk blinded to the case or control status of participants? (12) Were key potential confounding variables measured and adjusted statistically in the analyses? If matching was used, did the investigators account for matching during study analysis?

3

Indicators to facilitate the early identification of patients with major depressive disorder in need of highly specialized care: A concept mapping study

van Krugten FCW
Goorden M
van Balkom AJLM
Spijker J
Brouwer WBF
Hakkaart-van Roijen L

Depression and Anxiety. 2018;35(4):346-352.

Abstract

Objectives

Early identification of the subgroup of patients with major depressive disorder (MDD) in need of highly specialized care could enhance personalized intervention. This, in turn, may reduce the number of treatment steps needed to achieve and sustain an adequate treatment response. The aim of this study was to identify patient-related indicators that could facilitate the early identification of the subgroup of patients with MDD in need of highly specialized care.

Methods

Initial patient indicators were derived from a systematic review. Subsequently, a structured conceptualization methodology known as concept mapping was employed to complement the initial list of indicators by clinical expertise and develop a consensus-based conceptual framework. Subject-matter experts were invited to participate in the subsequent steps (brainstorming, sorting and rating) of the concept mapping process. A final concept map solution was generated using non-metric multidimensional scaling and agglomerative hierarchical cluster analyses.

Results

In total, 67 subject-matter experts participated in the concept mapping process. The final concept map revealed ten major clusters of indicators: depression severity, onset and (treatment) course, comorbid personality disorder, comorbid substance use disorder, other psychiatric comorbidity, somatic comorbidity, maladaptive coping, childhood trauma, social factors, and psychosocial dysfunction.

Conclusions

The study findings highlight the need for a comprehensive assessment of patient indicators in determining the need for highly specialized care, and suggest that the treatment allocation of patients with MDD to highly specialized mental healthcare settings should be guided by the assessment of clinical and non-clinical patient factors.

Introduction

Timely selection of the best initial treatment for patients with major depressive disorder (MDD) is critical to the goal of improving remission rates [92]. The often-applied stepped care approach in which patients indiscriminately receive brief and low-intensity intervention at start of treatment and intensifying efforts in case of insufficient signs of recovery, may, however, prevent the accurate and timely selection of the best initial treatment. Although the stepped care approach is considered a resource efficient approach for patients who recover with minimal intervention [93, 94], the effectiveness of this approach is questionable in patients who need subsequent referral to highly specialized mental healthcare services. Secondary or even tertiary referral to highly specialized mental healthcare services delays the initiation of appropriate treatment, which, in turn, is associated with poor treatment outcomes in terms of relapse, recurrence and chronicity [29, 30, 95]. An alternative to the stepped care approach is matched care. In this approach, patient management and initial treatment allocation is tailored to the individual patient needs [96, 97]. Successful application of this approach may reduce the number of treatment steps needed to achieve and sustain an adequate treatment response, benefit the quality of life of patients and increase the cost-effective use of resources.

A major problem in the application of matched care approach is the lack of clear individual patient indicators with which to match patients to the available treatment settings. In recent years, a wide array of individual patient factors has been examined to inform initial treatment selection in patients with MDD [98-102]. Despite some progress, these have thus far not demonstrated their value in clinical practice and some approaches like neuroimaging are not feasible for use in daily clinical practice [103]. Information of individual patient factors collected as part of routine assessment procedures in the diagnostic phase after referral, however, has the potential to aid clinicians in the early identification of the patients with MDD in need of highly specialized care. The aim of this study, therefore, was to identify a range of clinical and non-clinical factors of patients with MDD in need of highly specialized care that could serve as input for the development of a decision support algorithm.

Methods

Prior to the study period, a small working group was formed comprising thirteen leading Dutch experts in the field of MDD from nine mental healthcare institutions. The workgroup included academically affiliated and community-based practicing MDD

specialists. This study progressed through two primary phases. First, a systematic review of the literature of the PubMed and the PsycINFO databases following PRISMA guidelines [65] was conducted to serve as a scientific foundation. The aim of this systematic review was to comprehensively cover the factors associated with a broad range of unfavourable clinical outcomes in patients with MDD for which more intensive treatment is indicated. The systematic search of all databases yielded a total of 7,360 references, of which sixteen were eligible for inclusion. Based on the included papers, an initial list of 48 indicators of patients with a depression in need of highly specialized care was generated (see [104; Chapter 2] for details of this review). Subsequently, a structured conceptualization methodology known as concept mapping [105] was employed to complement the initial list of indicators by clinical expertise and develop a consensus-based conceptual framework. Concept mapping is a method that integrates a qualitative research design with quantitative analytic techniques to conceptualize a phenomenon of interest [105, 106] and has been used in a wide variety of studies, including measurement development [107-110]. In general, the concept mapping process involves the following five steps: (1) preparation; (2) brainstorming; (3) sorting and rating; (4) statistical analysis; and (5) interpretation [105, 111]. These steps are described below, along with details of how we implemented them in this study.

Step 1. Preparation

During the first step of concept mapping, a focal question was developed and relevant subject-matter experts were selected. In collaboration with the small working group of experts, we developed a focal question for item elicitation. Our focal question was: "Which criteria distinguish depressive patients in need of a highly specialized mental healthcare treatment from patients in need of regular specialized mental healthcare treatment?". The focal question was developed to elicit a list of participants' ideas that were then analysed for the study.

Working group members identified and selected subject-matter experts from a broad range of disciplines. These experts were identified and selected based on their expertise in the assessment and/or treatment of patients with MDD or involvement in MDD research. In total, 184 national and international experts were invited to participate in the subsequent data collection activities. At the time of data collection, all participants were asked to sign an electronic consent form for participation and complete a brief demographic questionnaire. All data collection activities for this study were performed in English and Dutch in order to facilitate national and international subject-matter expert participation.

Step 2. Brainstorming

In step 2, working group members distilled the original 48 indicators into a list with distinct statements by eliminating duplicate statements, editing statements for clarity, or combining similar statements. This process resulted in a list of 38 mutually exclusive indicators of patients with MDD in need of highly specialized care. Participants were then asked to individually review this list of indicators and engage in a brainstorming session to generate additional indicators. Brainstorming took place through a web-based system specifically designed for concept mapping (Concept Systems® software, Incorporated, Ithaca, New York). The open-ended focal question mentioned under step 1 was used to elicit criteria from participants. In response to the focal question, participants were asked to generate as many criteria as possible and enter them into the system. The participants had four weeks to respond to our request. During this four-week period, they had the option of entering criteria in more than one session.

Step 3. Sorting and rating

Following procedures recommended by Trochim [112], sorting and rating activities were performed as an individual activity via the aforementioned web-based program. Participants were asked to sort the criteria into categories based on the principle of similarity, thereby building thematically related sets of items. Specifically, participants were instructed to group criteria in a way that 'made sense to them' and label their final groupings accordingly. The sole restrictions were: (1) all criteria cannot be placed into a single category; (2) a criterion cannot be placed simultaneously into two separate categories; (3) categories named 'Miscellaneous' and 'Other' that group together dissimilar statements are not allowed; and (4) criteria cannot be sorted according to priority or value, such as 'Important', or 'Hard to do'.

After completing the sorting activity, participants were asked to rate each individual criterion on how important it was to distinguish between patients in need of highly specialized care from patients in need of specialized care. Responses were recorded on a 5-point Likert scale ranging from 1 (not important at all) to 5 (extremely important).

Step 4. Statistical analysis

Concept Systems software was used to analyse the data generated from the sorting and rating exercise. Three statistical procedures were sequentially performed. First, a non-metric multidimensional scaling (MDS) analysis was carried out to plot the criteria and their cohesion on a two-dimensional plane. The analysis yielded a so-called "point map" on which the proximity of the points represents the frequency with which the criteria were sorted together by each of the individual participants. Points located closer to each other on the point map represent criteria sorted together most often, whereas

points located further apart represent criteria sorted together less frequently. A stress value was calculated as part of the multidimensional scaling analysis to indicate how well the two-dimensional configuration maps the original data. The stress value is an index of the goodness of fit of the MDS solution and ranges from 0 to 1, with lower values indicating a better fit. Subsequently, an agglomerative hierarchical cluster analysis using Ward's minimum variance method [113] was carried out to partition the resulting MDS configuration into non-overlapping clusters, thereby creating initial cluster maps. Mean importance ratings of the clusters were computed by averaging the average rating of each criterion in the clusters. Finally, paired t-tests were carried out to compare the mean importance ratings of the various clusters. To adjust for multiple testing, a Bonferroni correction was used, dividing the conventional alpha of 0.05 by the number of independent tests.

Step 5. Interpretation

Since there is no objective standard or mathematical solution through which a final number of clusters can be selected [114], working group members discussed the preliminary cluster solutions from the hierarchical cluster analyses to reach consensus on the optimal cluster number for answering the focal question. Following recommendations by Kane and Trochim [111], a range of cluster solutions was examined in a reverse stepwise cluster-reduction process. In this process, two clusters merge (e.g. from 14 to 13 clusters) at each reverse step. Working group members worked backwards from 20 clusters and examined successively lower cluster solutions. At each level, a judgment was made about whether the merger made conceptual and interpretive sense until a cluster level was reached that yielded the fewest number of clusters but still retained the maximum amount of substantive information. In a digital survey, working group members were then asked to review the within-cluster coherence of content and suggest criteria that could be moved from one cluster to another to increase conceptual clarity and assign cluster-labels to the resulting clusters. Informed by the gathered working group input, the clusters were assigned final labels and some criteria were reallocated to a conceptually more appropriate cluster.

Results

Expert participation

In total, 67 out of the 184 invited subject-matter experts participated in one or more of the steps of the concept mapping process. The mean age of the experts was 50.42 years (SD=10.93) and 41.54% (n=27) were female. The mean years of work experience in the assessment and/or treatment of patients with MDD or involvement in MDD

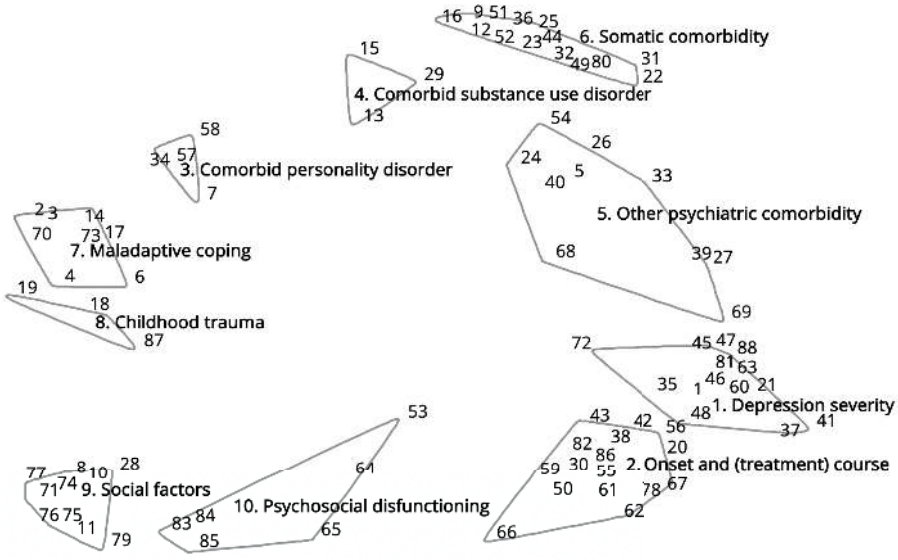
research was 23.31 years (SD=11.30). The majority of the experts were psychiatrists (n=44, 67.7%), followed by clinical psychologists (n=12, 18.5%), clinical researchers (n=4, 6.2%), psychotherapists (n=4, 6.2%), and physicians (n=1, 1.5%). There was equal representation of experts working in specialized mental healthcare settings and highly specialized mental healthcare settings (n=34, 52.3% and n=31, 47.7% respectively).

Concept mapping results

A total of 50 items were generated during the brainstorming stage and added to the initial list of 38 indicators derived from the systematic review, resulting in a list of 88 putatively relevant indicators of patients with MDD in need of highly specialized care. This list of 88 items was used in the subsequent sorting and rating steps of the concept mapping process.

Forty-three (n=43) experts sorted the 88 indicators into an average of 9.00 piles (SD=3.04). The stress value of the MDS-solution was 0.23, falling within the average range (0.15 to 0.35) of stress values typically attained in concept mapping studies [112]. The ten-cluster concept map solution produced by the participant sorts and subsequent analysis is presented in Figure 3.1. This cluster solution provided the maximum number of interpretable clusters without losing distinctions between groups of indicators. The numbers on the map correspond to the indicators listed in Appendix 3.1. Numbers closer together represent indicators that were more frequently sorted together than were indicators represented by points further apart. The more distance between numbers the less often they were sorted together (i.e., the less conceptually similar they were viewed by participants). Each cluster consists of indicators that were sorted together more frequently and contribute to an overarching conceptual domain. The shape and size of the clusters reflect the breadth or specificity of the clusters, with large clusters typically covering a broader, less well-defined concept than smaller clusters. The ten clusters were labelled: (1) depression severity; (2) onset and (treatment) course; (3) comorbid personality disorder; (4) comorbid substance use disorder; (5) other psychiatric comorbidity; (6) somatic comorbidity; (7) maladaptive coping; (8) childhood trauma; (9) social factors; and (10) psychosocial dysfunction. The overarching conceptual domains, sample indicators per conceptual domain, mean cluster ratings and cluster rankings are presented in Table 3.1. Mean importance ratings (i.e., ratings averaged across all indicators within a cluster) ranged between 2.53 and 4.42. On average, items in the depression severity cluster were rated most important to distinguish between patients in need of highly specialized care from patients in need of specialized care (M=4.42), followed by items in the psychiatric comorbidity cluster (M=4.18), and somatic comorbidity cluster (M=3.95). No consistent significant differences were found between mean importance ratings of the clusters.

Figure 3.1. Concept map of the main indicators of patients with MDD in need of highly specialized care (stress value = 0.23).



Clusters represent the overarching conceptual domains of the 88 indicators of patients with MDD in need of highly specialized care. Labels were suggested by working group members and finalized by the project team. Numbers correspond to the indicators that were sorted into each category. Indicators that are closer together indicate higher degrees of similarity based on sorting.

Discussion

The aim of this study was to identify patient-related indicators that could facilitate the early identification of the subgroup of patients with MDD in need of highly specialized care. Drawing on clinical expertise and a literature review, a concept mapping approach was employed to develop a consensus-based conceptual framework. Concept mapping is a mixed-method participatory approach that facilitated the delineation of a shared understanding of clinical and non-clinical patient indicators that may justify referral to highly specialized mental healthcare programs. In total, 88 putatively relevant indicators of patients with MDD in need of highly specialized care were generated and categorized into the following ten overarching conceptual domains: depression severity, onset and (treatment) course, comorbid personality disorder, comorbid substance use disorder, other psychiatric comorbidity, somatic comorbidity, maladaptive coping, childhood trauma, social factors, and psychosocial dysfunction.

Table 3.1. Conceptual domains, sample indicators, importance rating and ranking for the ten clusters.

Cluster		Sample indicators[a]		Importance	
				Mean rating[b]	Ranking
1	Depression severity	–	Greater depressive symptom severity	4.42	1
		–	Psychotic symptoms		
		–	Current suicidal risk		
		–	Higher rates of melancholic features		
2	Onset and (treatment) course	–	Younger age of onset	3.80	4
		–	Longer duration of index depressive episode		
		–	More lifetime episodes		
		–	Lack of remission or partial remission after the previous depressive episode		
3	Comorbid personality disorder	–	Higher Axis II personality pathology score	3.68	6
		–	Comorbid personality disorder		
4	Comorbid substance use disorder	–	Alcohol abuse	2.86	9
		–	Substance abuse		
5	Other psychiatric comorbidity	–	A higher number of comorbid psychiatric disorders	4.18	2
		–	Comorbidity with ADHD		
		–	Comorbidity with OCD		
		–	Comorbid (generalized) anxiety disorder		
6	Somatic comorbidity	–	Greater levels of general medical comorbidity	3.95	3
		–	Worse physical health function		
		–	A higher number of pain locations		
		–	Lower physical quality of life		
7	Maladaptive coping	–	Disadaptive coping	2.96	8
		–	High external locus of control		
		–	Less positive outcome expectancies		
8	Childhood trauma	–	Higher prevalence of childhood trauma	3.69	5
		–	Higher levels of trauma sequelae		
9	Social factors	–	No social support	2.53	10
		–	Lower monthly household income		
		–	Unemployment		
10	Psychosocial dysfunction	–	Worse social functioning	3.21	7
		–	Worse work function and social adjustment		
		–	More impaired daily function		

ADHD = Attention-Deficit/Hyperactivity Disorder; OCD = Obsessive-Compulsive Disorder.

[a] Indicators for which there was the most consensus among participants regarding the categorization within the cluster.

[b] Importance was rated on a 5-point Likert scale ranging from 1 to 5, with higher scores reflecting greater importance to distinguish between patients in need of highly specialized care and patients in need of specialized care.

To our knowledge, this is the first time that indicators of patients with MDD in need of highly specialized care emerging from the literature are appraised, refined, and complemented by clinical expertise. The resulting overarching conceptual domains of this concept mapping study repeat, to a certain extent, the main indicators of patients with MDD in need of highly specialized care found in the literature review (see [104; Chapter 2] for details of the review). Of the 88 putatively relevant indicators, 38 had

been identified in the literature review but were made more detailed, worded more precisely, and complemented by clinical expertise before being used in the subsequent sorting and rating steps of the concept mapping process. As such, the use of clinical expertise in addition to evidence from the literature, allowed the summarization of patient indicators emerging from the literature in well-defined overarching domains. These domains can serve as a starting point for the development of a selection algorithm, which, in turn, may contribute to systematic, evidence-based treatment selection in patients with MDD.

At the domain level, importance ratings ranged from 2.53 to 4.42 on a 5-point Likert scale ranging from 1 (not important at all) to 5 (extremely important). Domains with relatively low mean importance ratings appear to cover the non-clinical patient indicators such as treatment-interfering maladaptive coping (domain 7) and social factors maintaining the depression (domain 9), whereas domains of relatively higher importance seem to describe the clinical patient indicators such as depression severity (domain 1), psychiatric and somatic comorbidity (domains 2, 6 and 3) and childhood trauma (domain 8). Although the high mean importance ratings of domains covering clinical patient indicators is consistent with findings indicating that most clinical decisions are largely based on 'traditional' clinical patient factors [116], the impact of each domain on referral decisions in patients with MDD remains to be validated in an observational study. Future research should examine the relative importance and possible synergy of action between the domains.

This study has a number of strengths, including the systematic step-by-step procedure of the concept elicitation procedure, the relatively high number of participants, and the use of clinical expertise in addition to evidence from the literature. The present results should, however, also be viewed in the light of some limitations of this study. First, aiming for the early identification of patients with a highly specialized care need and the timely allocation of those patients to highly specialized mental healthcare settings, presupposes that there is something like a 'right place' and that getting there sooner is better than later. Although highly specialized care has been demonstrated to improve clinical outcomes in patients with complex and severe conditions in other areas of medicine [87], the net benefit of highly specialized care in patients with MDD has, however, not yet been demonstrated. Future studies should therefore address the evaluation of the impact of highly specialized care on patient outcomes in this population. Second, in line with the inclusion criteria of the systematic review, the study results are restricted to patients aged 18 and over with a primary diagnosis of MDD treated in psychiatric specialized and highly specialized outpatient clinics. Hence, the findings of this study cannot be generalized to non-clinical samples, children and

adolescents. Third, although the number of subject-matter experts that participated in one or more of the steps of the concept mapping process falls within the average range (20 to 649) of participants in concept mapping research [117], it is unclear whether the participants' conceptualization is representative of the larger population. In addition, although effort was made to include subject-matter experts from a broad range of disciplines and countries, the majority of the participants were psychiatrists and worked as treating clinicians and/or researchers in the Netherlands. A larger and more heterogeneous sample of the population might have resulted in a broader range of perspectives and enhanced the generalizability of the findings. Fourth, although involvement of experts is in accordance with evidence-based medicine [118], the patient indicators generated by the clinicians may be biased by pre-existing perceptions, beliefs or attitudes. Future research using a larger and more heterogeneous sample should explore to what extent the results are valid, stable and generalizable.

Despite these limitations, the results of the present study provide a practical first step towards the early identification of patients with MDD in need of highly specialized care. The study findings highlight the need for a comprehensive assessment of patient indicators in determining the need for highly specialized care, and suggest that the treatment allocation of patients with MDD to highly specialized mental healthcare settings should be guided by the assessment of clinical and non-clinical patient indicators. The results of this study can serve as input for the development of a decision support algorithm to aid clinicians in the treatment allocation of patients with MDD in need of highly specialized care. Such an algorithm may be used to objectify clinical impressions and ultimately assist clinicians in selecting the most appropriate treatment strategy in a given clinical situation. As such, the results of this study have the potential to support and enhance personalized medicine, in which patient management and treatment is tailored to the individual patient needs [119]. Additional research is needed to evaluate the relative importance and possible synergy of action between the identified patient factors and the selection of an optimal decision threshold to distinguish patients with and without a need for highly specialized MDD care.

Appendix 3.1 – Cluster indicators and average cluster ratings

Table 3.1.1. Cluster indicators and average cluster ratings.

Cluster	Indicator	Average Rating	Systematic review[a]	Brainstorm[b]
1. Depression severity		**4.42**		
1.	Atypic presentation / course			X
21.	Severe suicidality			X
35.	Severe depression			X
37.	Need for ECT			X
41.	Need for specialized care (e.g., ECT)			X
45.	(Severe) psychomotor retardation			X
46.	(Severe) nihilism			X
47.	Psychotic symptoms			X
48.	Greater depressive symptom severity		X	
60.	Current suicidal risk		X	
63.	Higher rates of melancholic features		X	
81.	An increased likelihood of atypical symptom features			X
88.	Features of catatonia like repetitive movement abnormalities			X
72.	Higher levels of rumination		X	
2. Onset and treatment course		**3.80**		
20.	Nonresponse to first antidepressant treatment lifetime		X	
30.	High recurrence rate			X
38.	Chronic depression			X
42.	Worsening clinical course			X
43.	Treatment resistance			
50.	Younger age of onset		X	
55.	Longer duration of index depressive episode			X
56.	History of prior suicide attempts		X	
59.	More lifetime episodes		X	
61.	Longer time since first onset		X	
62.	Fewer prior episodes of depression		X	
66.	Less likely to meet criteria for chronic depression		X	
67.	Higher trial numbers of antidepressants			X
78.	Shorter current episode		X	
82.	More hospitalizations		X	
86.	Lack of remission or partial remission after the previous depressive episode		X	

Table 3.1.1. Cluster indicators and average cluster ratings. (continued)

Cluster	Indicator	Average Rating	Systematic review[a]	Brainstorm[b]
3. Comorbid personality disorder		**3.68**		
7.	Attachment disorder			X
34.	Personality disorder			X
57.	Higher Axis II personality pathology score		X	
58.	Comorbid personality disorder			X
4. Comorbid substance use disorder		**2.86**		
13.	Alcohol abuse			X
15.	Drug use and addiction			X
29.	Substance abuse			X
5. Other psychiatric comorbidity		**4.18**		
5.	Comorbidity with ADHD			X
24.	Comorbidity with OCD			X
26.	Obsessive repetition of one somatic complaint			X
33.	Cognitive impairment			X
40.	Severe comorbidity			X
54.	Comorbid (generalized) anxiety disorder		X	
68.	Greater number of lifetime comorbid psychiatric disorders			X
27.	A higher number of comorbid psychiatric disorders		X	
69.	More symptoms of bipolarity		X	
39.	Higher levels of chronic PTSD		X	
6. Somatic comorbidity		**3.95**		
9.	Severe neck, chest and abdominal pain		X	
12.	Lower physical quality of life		X	
16.	Obesity			X
23.	Severe somatic comorbitity			X
25.	Comorbidity with newly arised neurologic signs			X
32.	Presence neurodegenerative disease like parkinson disease's			X
36.	Comorbidity with chronic medical illness			X
44.	Somatic comorbidity that interferes with antidepressive treatment			X
51.	Greater levels of general medical comorbidity		X	
52.	Worse physical health function		X	
80.	A higher number of pain locations		X	
49.	Higher severity of pain		X	
22.	Need for somatic interventions			X
31.	Polypharmacy (esp. with somatic medications)			X

Table 3.1.1. Cluster indicators and average cluster ratings. (continued)

Cluster	Indicator	Average Rating	Systematic review[a]	Brainstorm[b]
7. Maladaptive coping		**2.96**		
2.	Passive coping			X
3.	No understanding that change is something clients have to do themselves			X
14.	Disadaptive coping			X
73.	High external locus of control			X
4.	Lack of psychological perspective on symptoms			X
6.	Low therapy compliance			X
17.	Low self esteem			X
70.	Less positive outcome expectancies			X
8. Childhood trauma		**3.69**		
18.	Higher levels of trauma sequelae		X	
19.	Greater levels of chilhood emotional abuse		X	
87.	Higher prevalence of childhood trauma		X	
9. Social factors		**2.53**		
8.	No social support			X
28.	Social isolation			X
71.	No partner			X
74.	Less education		X	
75.	Lower monthly household income		X	
76.	Unemployment		X	
77.	Non-Western origin		X	
79.	Higher educational level		X	
11.	Older age		X	
10.	Prior job loss		X	
10. Psychosocial dysfunction		**3.21**		
83.	Impaired psychosocial functioning		X	
84.	Worse social functioning			X
85.	More time absent from work			X
64.	Lower quality of life		X	
65.	Worse work function and social adjustment		X	
53.	More impaired daily function			X
Total			**38**	**50**

[a] Indicators identified in the systematic review (see [104; Chapter 2] for details of this review).
[b] Indicators generated in the brainstorming stage of the concept mapping procedure.

4

The Decision Tool Unipolar Depression (DTUD): A new measure to facilitate the early identification of patients with major depressive disorder in need of highly specialized care

van Krugten FCW
Goorden M
van Balkom AJLM
van Oppen P
Ruhé HG
van Schaik DJF
Brouwer WBF
Hakkaart-van Roijen L

BMC psychiatry. 2019;19(1):179.

Abstract

Objectives

Selection of the optimal initial treatment in patients with major depressive disorder (MDD) in need of highly specialized care has the potential to benefit treatment outcomes and cost-effectiveness of treatment strategies. However, to date, there is a paucity of measures that could guide the selection of the initial treatment, in particular to indicate which patients with MDD are in need of highly specialized care. Recognizing this gap, this paper reports on the development and psychometric evaluation of the Decision Tool Unipolar Depression (DTUD), aimed to facilitate the early identification of patients with MDD in need of highly specialized care.

Methods

The DTUD was developed using a mixed-method approach, consisting of a systematic review and a concept mapping study. To evaluate the psychometric features of the DTUD, a cross-sectional multicenter study was conducted. A total of 243 patients with MDD were evaluated with the DTUD. Feasibility was operationalized as the time required to complete the DTUD and the content clarity of the DTUD. Inter-rater reliability was evaluated using Krippendorf's alpha. The Maudsley Staging Method (MSM) and the Dutch Measure for quantification of Treatment Resistance in Depression (DM-TRD) were administered to assess the convergent validity. A receiver operator characteristic curve was generated to evaluate the criterion validity and establish the optimal cut-off value of the DTUD.

Results

The mean administration time was 4.49 min (SD=2.71), and the content of the total DTUD was judged as clear in 94.7% of the evaluations. Inter-rater reliability values ranged from 0.69 to 0.91. Higher scores on the DTUD were associated with higher scores on the MSM (r_s=0.47) and DM-TRD (r_s=0.53). Based on the maximum Youden index (0.494), maximum discrimination was reached at a cut-off score of ≥5 (sensitivity 67%, specificity 83%).

Conclusions

The DTUD demonstrated to be a tool with solid psychometric properties and, therefore, is a promising measure for the early identification of patients with MDD in need of highly specialized care. Use of the DTUD has the potential to facilitate the selection and initiation of the optimal initial treatment in patients with MDD, which in turn may improve the clinical effectiveness and cost-effectiveness of treatment strategies.

Introduction

Since delayed initiation of appropriate treatment in patients with major depressive disorder (MDD) has been associated with relapse, recurrence and chronicity [30, 120, 121], early initiation of the optimal type and intensity of intervention is considered essential [122]. The stepped care model of healthcare delivery, according to which many parts of healthcare systems are organized and sometimes incentivized to work [123], may however delay the initiation of the optimal type and intensity of intervention. Within the stepped care approach, patients first receive the briefest, least intrusive, or least costly intervention, and only 'step up' the treatment pathway in case of changing treatment needs or insufficient health gains from initial treatment [27]. Although the stepped care model of healthcare delivery is considered an appropriate approach in patients who recover with low intensity treatments [93, 94], the effectiveness and cost-effectiveness of the stepped care model is questionable in patients who, identifiably, are in need of high intensity treatment [27]. Subsequent referral of these patients to highly specialized mental healthcare (i.e. tertiary mental healthcare) is likely to prolong the treatment course and compromise clinical and functional outcomes and cost-effectiveness of treatments. Selection of the optimal initial treatment in patients with MDD in need of highly specialized care is therefore warranted, as it can improve the effectiveness and cost-effectiveness of treatment paths, but strongly relies on the availability of psychometrically sound instruments to aid clinicians in the early identification of these patients [122, 124].

Several measures are available to screen for MDD and assess its severity in clinical and research settings [125-127]. However, to date, there is a paucity of measures that facilitate the selection of the optimal initial treatment, in particular to indicate which patients with MDD are in need of highly specialized care. Recognizing this gap, in this paper we report on the development and psychometric evaluation of the Decision support Tool for the assessment of highly specialized mental healthcare needs of patients with a Unipolar Depression, or the "Decision Tool Unipolar Depression" (DTUD) for short. The DTUD is a ten-item clinician-administered instrument to facilitate the early identification of patients with MDD in need of highly specialized care. The focus of this paper is on describing the development of the DTUD and presenting the first results regarding its feasibility, inter-rater reliability, convergent validity, and criterion validity.

Methods

Definition of terms

As illustrated by the World Health Organization's Optimal Mix of Services Pyramid [23], most people with mental health problems are ideally treated in primary care services. When the mental health needs require intervention beyond that which can be provided by primary care services, the patient should be referred to specialized mental healthcare services (i.e. secondary mental healthcare) [23]. Specialized mental healthcare includes the mental health services provided in community mental health centers and general hospitals [23]. Highly specialized mental healthcare (i.e. tertiary mental healthcare) includes specialized interventions provided by highly-trained mental healthcare professionals with expertise in a given area to patients with mental health problems that require intervention over and above those provided in specialized mental healthcare [24, 25]. Given the required level of staff expertise, management, security, and resources of highly specialized mental healthcare, those services are frequently, but not necessarily, affiliated with academic medical healthcare centers [25].

Decision Tool Unipolar Depression (DTUD) development

Aim of the development of the DTUD was to create a valid and reliable, yet at the same time short and easy to score clinician-administered measure to facilitate the early identification of the subgroup of patients with MDD who are in need of highly specialized mental healthcare. The development of the DTUD comprised the following three phases: (1) identification of indicators of patients with MDD in need of highly specialized care through a systematic literature review; (2) development of a conceptual framework to inform item generation; and (3) development of the measure and evaluation of face validity and feasibility.

In the first phase of the development of the DTUD, a systematic literature review was carried out to provide a scientific foundation for the selection of items included in the resultant measure [104; Chapter 2]. The PubMed and PsycINFO electronic databases were searched for studies published between January 2000 and January 2015 reporting indicators of patients with MDD in need of highly specialized care. The search retrieved 7,360 references, of which sixteen met the inclusion criteria. Two reviewers determined study eligibility, reviewed study quality, and performed data abstraction. From the included studies, 48 indicators of patients with MDD in need of highly specialized care were abstracted. For more details on the systematic review we refer to Van Krugten et al. [104; Chapter 2].

In the second phase of the development of the DTUD, concept mapping methodology [105] was used to generate a conceptual framework to guide tool development [115; Chapter 3]. In total, 67 MDD experts participated in the subsequent steps of the concept mapping process. During the first step of the concept mapping process (i.e. the brainstorming step), participating experts were asked to review the indicators from the systematic literature review, and, when necessary, add additional indicators that could discriminate MDD patients with and without a highly specialized care need. In the second step of the concept mapping process (i.e. the sorting step), participants individually sorted the resulting indicators from the brainstorming step into conceptual groupings. The data from the sorting step were analyzed using nonmetric multidimensional scaling and agglomerative hierarchical cluster analyses, resulting in a ten-cluster concept map solution. In a consensus meeting, consortium members reviewed the concept map and assigned labels to each of the ten clusters. The ten clusters (i.e. overarching domains) of indicators of patients with MDD in need of highly specialized care were assigned the following labels: depression severity, onset and (treatment) course, comorbid personality disorder, comorbid substance use disorder, other psychiatric comorbidity, somatic comorbidity, maladaptive coping, childhood trauma, social factors, and psychosocial dysfunction. For more details on the concept mapping study we refer to Van Krugten et al. [115; Chapter 3].

In the third phase of the development of the DTUD, members of Decision Tool Unipolar Depression Consortium generated the draft DTUD based on the resulting overarching domains from the concept mapping study (phase 2). In a consensus meeting, each of the overarching domains was operationalized as a dichotomous item. In order to evaluate the feasibility and face validity of the DTUD, the draft version of the DTUD was pilot-tested in a convenience sample of 46 patients aged 18 years or older with a (principal) primary diagnosis of MDD referred for treatment to a specialized or highly specialized treatment center in the Netherlands. Participating clinicians were asked to complete a web-based survey comprising the draft version of the DTUD, comment on the clarity of content of the DTUD and register the time needed to complete the DTUD. In a three-hour consensus meeting, consortium members reviewed the pilot data and made minor revisions to the wording of the draft version, resulting in the final version of the DTUD. The resulting DTUD is a ten-item clinician-administered instrument designed to facilitate the early identification of individual patients with MDD in need of highly specialized mental healthcare. Each item has two response options (“Yes” and “No”). The total score is calculated by summing the scores of the ten items and ranges from 0 to 10. The abbreviated items of the DTUD are listed in Table 4.1. An English translation of the complete DTUD can be found in Appendix 4.1.

Table 4.1. Items, response options and scoring system of the DTUD.

Item[a]		Response options	Score
1	Severe depression	Yes No	1 0
2	Previous unsuccessful treatment of the index depressive episode in specialized care *and* a recurrent *or* chronic course	Yes No	1 0
3	Treatment-interfering comorbid personality disorder	Yes No	1 0
4	Treatment-interfering comorbid substance use disorder	Yes No	1 0
5	Other treatment-interfering psychiatric comorbidity	Yes No	1 0
6	Treatment-interfering somatic comorbidity	Yes No	1 0
7	Treatment-interfering maladaptive coping	Yes No	1 0
8	Severe or longstanding childhood trauma	Yes No	1 0
9	Social factors maintaining the depression	Yes No	1 0
10	Severe psychosocial dysfunctioning	Yes No	1 0

[a] Item text is abbreviated. An English translation of the complete DTUD is presented in Appendix 4.1.

Study design and population

The aim of the present study was to evaluate the psychometric properties of the DTUD. To that end, a cross-sectional, observational multicenter study was carried out in six psychiatric specialized and highly specialized outpatient centers in The Netherlands. The Medical Ethical Committee of the Erasmus University Medical Center Rotterdam reviewed and approved the study (MEC-2015-670).

243 randomly selected outpatients referred for treatment of a current episode of MDD to one of the six participating sites were evaluated with the DTUD under routine care conditions. Study inclusion criteria were: aged 18 years or older and a primary (principal) diagnosis of MDD according to Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV criteria [59]. The DSM-IV axis I diagnosis was determined by the administration of a Dutch version of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) [129] or by a structured clinical interview using DSM-IV criteria.

Measures

In addition to the DTUD, the following instruments were administered:

- The *Maudsley Staging Method* (MSM) [130] is a five-item, clinician-administered instrument designed to quantify (future) treatment resistant depression (TRD). The MSM comprises the following three dimensions: duration, severity and failed treatments in the current episode of depression. The total score ranges from 3 to 15, and may be categorized into three staging categories: mild (3-6), moderate (7-10), and severe (11-15).
- The *Dutch Measure for quantification of Treatment Resistance in Depression* (DM-TRD) [131] is an eleven-item, clinician-administered instrument, and an extension of the MSM. In addition to the MSM dimensions, the DM-TRD comprises dimensions for functional impairment, comorbid anxiety and personality disorders and psychosocial stressors. The total score ranges from 2 to 27, with higher values indicating higher levels of TRD.

Procedures

Patients who were referred to one of the six participating clinics with a primary (principal) diagnosis of MDD were evaluated with the DTUD. Attending clinicians completed the DTUD at the end of the diagnostic phase, on the basis of the diagnostic results. In addition to the DTUD, the clinician administered the MSM and DM-TRD, recorded the patients' basic demographic information (age, sex), and answered two questions regarding the feasibility of the DTUD. The participating clinics entered the data in completely anonymized web-based case report forms as approved by the institutional review board.

Feasibility was operationalized as the time required to complete the DTUD, and the content clarity of the DTUD. Completion time was considered acceptable if the mean time taken to complete the DTUD was ≤10 minutes. The clarity of the total DTUD was scored with "Yes" or "No", and was considered acceptable if ≥90% of the informants evaluated the content of the DTUD as clear. Inter-rater reliability was assessed in a random subsample of 54 patients using pairs of independent ratings made by two clinicians present at the same admission interview. Assessment of the criterion validity of the DTUD was conducted in four out of six participating psychiatric clinics. Since a reference standard for the determination of need for highly specialized MDD care was not available, the experts' clinical judgement constituted the reference standard. At each clinic, two clinicians with extensive clinical experience in the treatment of depressive disorders, independently and blinded to the index score (i.e. DTUD), made a clinical judgment based on the patient's medical record as to whether the patient was in need of highly specialized care (Yes/No). An independent researcher verified the consistency between the two clinical judgments, and discrepancies were resolved by a consensus meeting with the first and second clinician.

Statistical analysis

All analyses were conducted using SPSS (Statistical Package for the Social Sciences) version 20.0 (IBM SPSS Version 20, IBM, New York, NY, USA). Statistical significance was inferred at $P<0.05$ (two-tailed). Demographic characteristics and feasibility outcomes were examined using descriptive statistics. Feasibility outcomes were evaluated according to the criteria outlined in the procedures section. Inter-rater reliability was assessed by Krippendorff's alpha for the individual items and total DTUD score [132, 133]. Krippendorff's alpha is a conservative reliability estimate for judgments made by any number of raters, and is adaptable to any level of measurement [134]. For each of the estimated Krippendorff's alpha values, 95% confidence intervals (CIs) were computed based on 10,000 bootstrap replications. Estimated Krippendorff's alpha values were evaluated against the minimum recommended reliability level of 0.667 [133]. Convergent validity was assessed by Spearman's correlation coefficients between total DTUD scale scores and total MSM and DM-TRD scores. Correlations of 0.10-0.30, 0.30-0.49 and >0.50 were considered as weak, moderate and strong, respectively [135]. The DTUD was hypothesized to have a positive correlation with the MSM and DM-TRD. A receiver-operating characteristic (ROC) curve was generated to assess the criterion validity of the DTUD. In order to determine the optimal cut-off score, a Youden index ($J = (\text{sensitivity}_c + \text{specificity}_c) - 1$) [136] was calculated for a range of cut-off scores. The cut-off score that corresponded to the highest Youden index was selected as the optimal cut-off score.

Results

Description of the study population

From November 2015 to April 2016, a total of 243 patients were studied. Table 4.2 summarizes the main demographic and clinical data of the patients. The mean age of the patients was 44.22 years (SD=12.64) and 60.49% (n=147) were female. The length of the index depressive episode was less than twelve months for 44.45%; one year to two years for 11.52%, and more than two years for 44.03% of the sample. Using DSM-IV specifiers, the majority of the patients were diagnosed with moderate (36.63%) or severe MDD without psychosis (34.98%). The mean total DTUD score was 3.70 (SD=2.00). Mean total MSM and DM-TRD scores were 6.71 (2.42) and 11.30 (3.67), respectively.

Feasibility

The mean administration time was 4.49 min (SD=2.71), and the content of the total DTUD was in 94.65% of the evaluations judged as clear. Two out of 48 clinicians sug-

Table 4.2. Demographic and clinical characteristics of the study sample.

	IRR sample[a]	**Criterion validity sample**[a]	**Total sample**
N	54	132	243
Age, years			
Mean (SD)	41.48 (12.15)	44.67 (11.89)	44.22 (12.64)
Range	23-66	22-69	18-78
Sex (n, %)			
Male	24 (44.44)	57 (43.18)	96 (39.51)
Female	30 (55.66)	75 (56.82)	147 (60.49)
Duration of current MDD episode (n, %)			
Acute (≤12 months)	27 (50.00)	56 (42.42)	108 (44.45)
Subacute (13-24 months)	7 (12.96)	13 (9.85)	28 (11.52)
Chronic (> 24 months)	20 (37.04)	63 (47.73)	107 (44.03)
Symptom severity of current MDD episode (n, %)			
Mild	14 (25.93)	24 (18.18)	48 (19.75)
Moderate	25 (46.30)	47 (35.61)	89 (36.63)
Severe without psychosis	11 (20.37)	49 (37.12)	85 (34.98)
Severe with psychosis	4 (7.41)	12 (9.09)	21 (8.64)
Total DTUD score			
Mean (SD)	3.85 (1.85)	3.65 (2.05)	3.70 (2.00)
Range	0.00-8.00	0.00-9.00	0.00-9.00
Total MSM score			
Mean (SD)	6.02 (2.16)	6.98 (2.42)	6.71 (2.42)
Range	3.00-13.00	3.00-13.00	3.00-13.00
Total DM-TRD score			
Mean (SD)	10.55 (3.13)	11.60 (3.97)	11.30 (3.67)
Range	6.00-23.50	3.00-23.50	3.00-23.50

DM-TRD = Dutch Measure for quantification of Treatment Resistance in Depression; DTUD = Decision Tool Unipolar Depression; IRR = Inter-Rater Reliability; MSM = Maudsley Staging Method; SD = Standard Deviation.
[a] Part of total sample.

gested the addition of a mid-point in the set of response options, such as "maybe" or "don't know". Three out of 48 clinicians expressed concern about the clarity of the items "social factors maintaining the depression" (item 9) and "severe psychosocial dysfunctioning" (item 10), and suggested the inclusion of examples and descriptions of both items to improve item clarity. Another suggestion included the addition of a statement according to which grade of diagnostic validity item 3 (comorbid personality disorder) should be determined - i.e. whether the item is met in case of a diagnosed personality disorder according to a structured interview such as the Structured Clinical Interview for DSM-IV (SCID) [129]), or also on the basis of a clinically suspected comorbid personality disorder, without administration of a formal structured interview.

Reliability

Inter-rater reliability was determined for 54 participants. As demonstrated in Table 4.3, the Krippendorf's alpha value of the total DTUD score was 0.82 (95% CI 0.76-0.87). The Krippendorff's alpha values of the individual items of the DTUD varied between 0.69 (95% CI 0.52-0.83) for comorbid personality disorder and 0.91 (95% CI 0.77-1.00) for comorbid substance use disorder. No item was below the minimum recommended reliability level of 0.667 [133].

Table 4.3. Krippendorff's alpha values of the DTUD (n=54, 95% CIs generated by 10,000 bootstrap replications).

Item		Krippendorff's alpha (95% CI)
1	Severity	0.81 (0.69-0.92)
2	Course	0.82 (0.68-0.92)
3	Comorbid personality disorder	0.69 (0.52-0.83)
4	Comorbid substance use disorder	0.91 (0.77-1.00)
5	Other psychiatric comorbidity	0.78 (0.64-0.90)
6	Somatic comorbidity	0.84 (0.64-0.92)
7	Coping	0.85 (0.74-0.94)
8	Childhood trauma	0.82 (0.70-0.92)
9	Social factors	0.78 (0.64-0.90)
10	Psychosocial functioning	0.73 (0.58-0.85)
Total DTUD score		**0.82 (0.76-0.87)**

CI = Confidence Interval; DTUD = Decision Tool Unipolar Depression.

Validity

As expected, higher scores on the DTUD were associated with higher scores on the MSM ($r_s(241)=0.47$ $P<0.001$) and DM-TRD ($r_s(241)=0.53$, $P<0.001$). Figure 4.1 and Table 4.4 summarize the operating characteristics of the DTUD. The area under the curve (AUC) was 0.81 (95% CI 0.73-0.87). Based on the maximum Youden index of 0.494, maximum discrimination was reached at a cut-off score of ≥5. This cut-off score demonstrated a sensitivity of 0.67 (95% CI 0.52-0.79) and a specificity of 0.83 (95% CI 0.73-0.90).

Discussion

This study evaluated the psychometric properties of the Decision Tool Unipolar Depression (DTUD) in the identification of patients with MDD in need of highly specialized mental healthcare. Overall, the results provide initial support for the psychometric properties of the DTUD. The DTUD demonstrated excellent feasibility and adequate inter-rater reliability. The associations with measures of TRD and health-related quality

Figure 4.1. ROC curve for the DTUD (area under the curve (AUC) = 0.81).

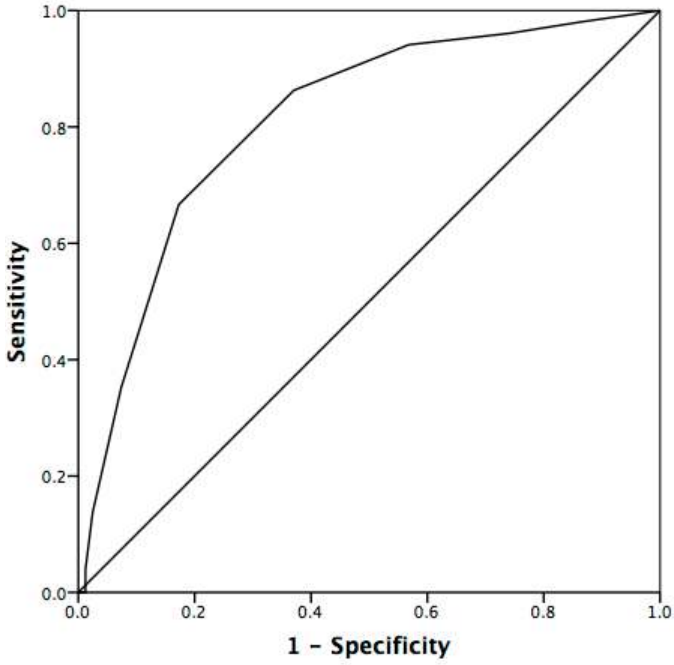


Table 4.4. Operating characteristics of the DTUD with the experts' clinical judgment constituting the criterion standard.

DTUD scale score	Sensitivity (95% CI)	Specificity (95% CI)	Youden index[a]
≥ 3	0.94 (0.84-0.99)	0.43 (0.32-0.55)	0.373
≥ 4	0.86 (0.74-0.94)	0.63 (0.52-0.73)	0.492
≥ 5	0.67 (0.52-0.79)	0.83 (0.73-0.90)	0.494
≥ 6	0.35 (0.22-0.50)	0.93 (0.85-0.97)	0.279

CI = Confidence Interval; DTUD = Decision Tool Unipolar Depression.
[a] Youden index = (sensitivity + specificity) - 1.

of life supported convergent validity. Furthermore, the DTUD demonstrated satisfactory criterion validity for use in clinical practice; a cut-off score of ≥5 was found to represent an optimal cut-off point for identifying patients with MDD in need of highly specialized care. The results support the use of the DTUD in busy, routine, outpatient specialized and highly specialized settings. Both the average completion time and content clarity of the questionnaire were within a-priori determined acceptability limits (≤10 minutes for completion time and ≥90% for clarity).

A noteworthy finding is that clinicians tend to disagree on the presence of a comorbid personality disorder. An analysis of the provided qualitative feedback regarding this item suggested that this may be due to the differential grade of diagnostic validity at which the presence of a comorbid personality disorder was determined (i.e. whether the item is met in case of a diagnosed personality disorder according to a structured interview, or also on the basis of a clinically suspected comorbid personality disorder, without the administration of a formal structured interview). Previous studies have shown that training on how to score an instrument can improve the reliability of a scale [137, 138]. Whether training also improves the reliability of the DTUD should be studied in future research.

The pattern of correlations between the DTUD and measures of (future) TRD and health-related quality of life supported convergent validity. Specifically, the DTUD was more strongly associated with the DM-TRD than with the MSM, suggesting that the MSM measures a more distantly related concept. This is to be expected since the DM-TRD is an extension of the MSM, additionally including items for functional impairment, comorbid anxiety, personality disorders and psychosocial stressors [131], all of which are well-known factors associated with unfavourable treatment outcome in MDD [139-144]. In addition, the DTUD showed good discriminative validity relative to the experts' clinical judgment of the need for highly specialized care (AUC=0.81). Based on the Youden index, maximum discrimination was reached at a cut-off score of ≥5, with a sensitivity of 67% and a specificity of 83%. A lower cut-off point (≥4) produced a similar Youden index value with higher sensitivity (86%) but at the cost of a lower specificity (63%). Given the limited capacity and higher costs of highly specialized services [23], higher specificity should be prioritized in order to decrease the rate of false positives, hence, a score of ≥5 is recommend and should be tested in future Decision Tool guided studies. For patients obtaining a DTUD score of 4, an initial evidence-based treatment in specialized mental healthcare should be combined with systematic monitoring and in case of inadequate treatment response, a quick, prioritized referral to highly specialized care should be strongly considered.

The key strengths of this study are the broad age-range of the sample, the extensive set of psychometric properties studied, and the nation-wide representation of the participating clinical sites (six clinics from across the country), which adds to the generalizability of the results. Further, to our knowledge, this is the first study in which a selection algorithm is developed and validated that facilitates the early identification of patients with MDD in need of highly specialized care. The results should, however, also be viewed in light of the study limitations. First, the feasibility of the DTUD was evaluated by completion time and content clarity; future studies could also assess

the feasibility of the DTUD with regard to item nonresponse. In the present study, an analysis of missing values was not possible since the web-based form was constructed in such a way that it required completion of all items. Second, the experts' clinical judgement constituted the reference standard for the evaluation of the criterion validity, which may have introduced subjective error. However, in the absence of a gold standard test for the identification or patients with MDD in need of highly specialized care, the experts' clinical judgement was considered the most adequate and clinically meaningful indication of highly specialized mental healthcare need. In addition, to reduce the subjective nature and increase the accuracy of the reference standard, the final clinical judgment was based on independent, dual examinations of comprehensive medical files by clinicians with extensive clinical experience in the treatment of depressive disorders. Third, the results reported in this paper represent a first examination of the DTUD psychometric properties. It was beyond the scope of this study to examine other issues, such as test-retest reliability, which should be examined in future studies. Fourth, it should be noted that the development of assessment tools typically requires a trade-off between feasibility (i.e. practicality) and validity (i.e. precision). Since the aim was to develop a simple, routine tool that is quick and easy to complete, the DTUD was constructed as a simple additive score of unweighted items. Future research might examine the relative importance of the individual items, as well the effect of the use of weighted items on the feasibility and validity of the DTUD. In addition, although the factors of the DTUD resulted as independent, distinct indicators of patients with MDD in need of highly specialized care from the concept mapping study [115; Chapter 3], there might be a potential for reduction of DTUD items through merging of potentially correlated items. Since the evaluation of the effect of merging potentially correlated items on the psychometric properties of the DTUD would require a new operationalization of items and subsequent psychometric testing, this evaluation should be addressed in future studies. Moreover, although the currently recommended cut-off value will likely generalize to similar psychiatric settings in The Netherlands, this remains to be validated. Finally, since the financing and organization of mental healthcare systems varies internationally [145, 146], future studies are needed to determine the appropriate cut-off value for other countries. In this regard, adapting the DTUD into other languages to test its suitability in similar groups of patients but in different healthcare systems may be beneficial to extend its cross-national robustness.

The results of the present study provide initial support for the psychometric properties of the DTUD. The DTUD proves to be a tool with excellent feasibility, adequate reliability and satisfactory validity and, therefore, is a promising instrument for the early identification of patients with MDD in need of highly specialized care. As such, the

results of this study have the potential to facilitate the selection and initiation of the optimal initial treatment in patients with MDD, which in turn may improve the clinical effectiveness and cost-effectiveness of treatment strategies.

Appendix 4.1 – Decision Tool Unipolar Depression (DTUD)

Figure 4.1.1. English translation of the full and final DTUD.

Decision Tool Unipolar Depression (DTUD) Version 1.0

Name of patient:

Patient medical record number:

1	Does the patient have severe depression? Indication: HDRS ≥25; IDS-C/SR ≥39; Q-IDS-C/SR ≥16; MADRS ≥31; BDI ≥30; CGI serious; DSM-IV 296.x3; PHQ ≥20	◯ yes ◯ no	Findings and/or level
2	In the index episode, have there been any unsuccessful previous treatments in specialised mental health care **and** does the patient have a - recurrent (more than 2 episodes in the past 5 years) **or** - chronic (>2 years) course of depression?	◯ yes ◯ no	Findings and/or level
3	Does the patient have a comorbid personality disorder according to DSM-IV/5 criteria that interferes with the depression treatment?	◯ yes ◯ no	Findings and/or level
4	Does the patient have a comorbid substance dependence disorder that interferes with the depression treatment?	◯ yes ◯ no	Findings and/or level
5	Does the patient have other severe psychiatric comorbidity that interferes with the depression treatment?	◯ yes ◯ no	Findings and/or level
6	Does the patient have somatic comorbidity that interferes with the depression treatment?	◯ yes ◯ no	Findings and/or level
7	Does the patient have a disadaptive coping style that interferes with the depression treatment? Hint: think of extreme avoidance or externalization	◯ yes ◯ no	Findings and/or level
8	Does the patient have a history of prolonged trauma/neglect in childhood?	◯ yes ◯ no	Findings and/or level
9	Are there any social factors contributing to the depression that are hard to influence?	◯ yes ◯ no	Findings and/or level
10	Does the patient exhibit severe psychosocial dysfunctioning?	◯ yes ◯ no	Findings and/or level

Total amount of positive (=yes) scores ≥5?

◯ yes ---> indicated for highly specialized care on the basis of the DTUD*

◯ no ---> not indicated for highly specialized care on the basis of the DTUD*

* This is a mental health care indication for adults. No rights can be derived.

5

Development and psychometric evaluation of the Decision Tool Anxiety Disorders, OCD and PTSD (DTAOP): Facilitating the early detection of patients with an anxiety disorder in need of highly specialized care

van Krugten FCW
Kaddouri M
Goorden M
van Balkom AJLM
Berretty EW
Cath DC
Hendriks GJ
Matthijssen SJMA
van Vliet IM
Brouwer WBF
Hakkaart-van Roijen L

PLOS ONE. 2021;16(8):e0256384.

Abstract

Objectives

Early identification of patients with an anxiety disorder, obsessive-compulsive disorder (OCD), or post-traumatic stress disorder (PTSD) in need of highly specialized care could facilitate the selection of the optimal initial treatment in these patients. This paper describes the development and psychometric evaluation of the Decision Tool Anxiety Disorders, OCD and PTSD (DTAOP), which aims to aid clinicians in the early identification of patients with an anxiety disorder, OCD, or PTSD in need of highly specialized mental healthcare.

Methods

A systematic literature review and a concept mapping procedure were carried out to inform the development of the DTAOP. To evaluate the psychometric properties of the DTAOP, a cross-sectional study in 454 patients with a DSM-IV-TR anxiety disorder was carried out. Feasibility was evaluated by the completion time and the content clarity of the DTAOP. Inter-rater reliability was assessed in a subsample of 87 patients. Spearman's rank correlation coefficients between the DTAOP and the EuroQol five-dimensional questionnaire (EQ-5D-5L) scores were computed to examine the convergent validity. Criterion validity was assessed against independent clinical judgments made by clinicians.

Results

The average time required to complete the eight-item DTAOP was 4.6 min and the total DTAOP was evaluated as clear in the majority (93%) of the evaluations. Krippendorff's alpha estimates ranged from 0.427 to 0.839. Based on the qualitative feedback, item wording and instructions were improved. As hypothesized, the DTAOP correlated negatively with EQ-5D-5L scores. The area under the curve was 0.826 and the cut-off score of ≥4 optimized sensitivity (70%) and specificity (71%).

Conclusions

The DTAOP demonstrated excellent feasibility and good validity, but weak inter-rater reliability. Based on the qualitative feedback and reliability estimates, revisions and refinements of the wording and instructions were made, resulting in the final version of the DTAOP.

Introduction

Although there is compelling evidence supporting the efficacy of psychological interventions in treating anxiety disorders, obsessive-compulsive disorders (OCD), and post-traumatic stress disorders (PTSD) [147-149], not all patients need and benefit from the same type and intensity of intervention [150]. In daily clinical practice, clinicians are faced with the challenge of providing the right treatment to the right patient at the right time and in the right place. The importance of this challenge is emphasised by the high demand for mental healthcare, relative to available resources, making it important to improve the cost-effectiveness of treatment decisions [23].

Although most patients with an anxiety disorder, OCD, or PTSD can and should be treated within primary care or secondary mental health services, a subset of patients requires additional expertise and support from highly specialized (i.e. tertiary) mental healthcare services [23]. Highly specialized mental health services are the services provided by highly trained mental health specialists to patients with complex mental health problems that cannot be fulfilled by primary and secondary mental healthcare services [24, 25]. Since delay in establishing the optimal treatment (intensity) has been associated with partial recovery and chronicity [151-153], early intensive treatment of patients who are predictively in need of highly specialized care is likely to reduce the treatment steps needed to achieve an adequate treatment benefit and prevent quality of life deterioration. This, in turn, may benefit the clinical and cost-effectiveness of treatments.

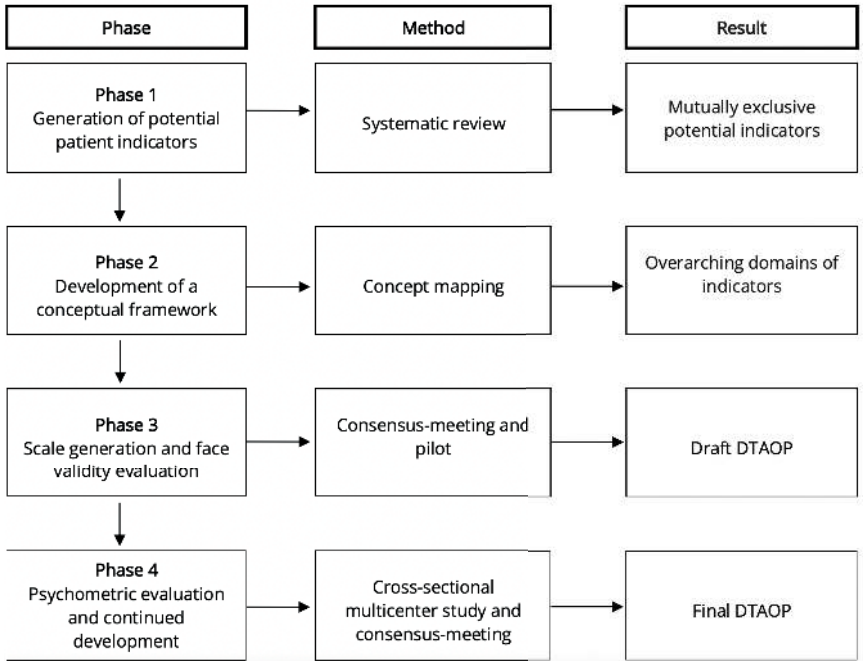
Although direct referral of patients with a severe and complex anxiety disorder, OCD, or PTSD to highly specialized mental health services may facilitate prompt, effective and efficient treatment, the effectiveness of this approach is highly dependent on the ability to identify these patients. Several measures to screen for anxiety disorders, OCD, and PTSD and assess their severity are available [154], yet psychometrically sound measures that aid the early identification of patients with a need for highly specialized care are lacking. Ideally, systematic and standardized pre-treatment assessments that not only capture the severity of the anxiety disorder, OCD, or PTSD itself, but also the complexity of the patient's overall clinical picture should be used to aid clinicians in matching the intensity of treatment to the individual patient needs [122]. Recognizing this gap, we report on the development and psychometric evaluation of the Decision Tool Anxiety Disorders, OCD and PTSD (DTAOP). The DTAOP is an eight-item clinician-administered instrument designed to enhance the systematic and standardized early identification of patients with an anxiety disorder, OCD, or PTSD in need of highly specialized care during the diagnostic phase after referral.

Methods

Phases of development

The DTAOP was initially designed for use in specialized (i.e. secondary) mental healthcare centers to inform referral decisions to highly specialized (i.e. tertiary) mental healthcare centers. The development process of the DTAOP consisted of the following four consecutive phases: (1) generation of potential indicators of patients with an anxiety disorder, OCD, or PTSD in need of highly specialized care; (2) development of a conceptual framework to guide item generation; (3) scale generation and assessment of face validity; and (4) evaluation of psychometric properties and continued development. See Figure 5.1 for a visual representation of the phases. Each phase was carried out in a manner consistent with previous research on Decision Tool development (see Van Krugten et al. [104; Chapter 2] and Van Krugten et al. [115; Chapter 3] for more details). See below for a summary of each of the four phases. The institutional Ethical Review Committee of the Erasmus University Medical Center Rotterdam, The Netherlands reviewed the study and declared that the Medical Research Involving Human Subjects Act (WMO) did not apply to this study and that therefore an official approval by the Ethical Review Committee was not required (MEC-2016-189). Written informed consent was obtained from all participants.

Figure 5.1. Development stages of the Decision Tool Anxiety Disorders, OCD and PTSD (DTAOP).



Throughout each phase of the development and psychometric evaluation process, clinicians and patients were consulted. Final decisions were made by the Decision Tool Anxiety Disorders, OCD and PTSD (DTAOP) consortium, comprising 14 Dutch clinicians in the field of anxiety disorders, OCD, and/or PTSD, 4 academics, and 1 patient representative.

Phase 1. Generation of potential patient indicators

As a first step in the development process, a systematic literature review was conducted to generate a list of potential indicators for the early detection of patients with an anxiety disorder, OCD, or PTSD likely in need of highly specialized care. Following the PRISMA guidelines [65], the PubMed and PsycINFO databases were searched for primary studies published between January 2000 and April 2015 reporting potential indicators of patients with an anxiety disorder, OCD, or PTSD in need of highly specialized care. In order to identify relevant search terms, 127 clinicians and 326 patients (aged 18-75 years) were invited to participate in a web-based survey; 99 and 231 participated respectively. In the survey, participants were asked to submit search terms to identify research articles reporting indicators of need for highly specialized care. Based on the submitted terms, search strategies were composed (see Appendix 5.1 for the search strategies). Two reviewers independently screened the titles and abstracts of all identified references, reviewed the full texts of potentially eligible articles, and performed data abstraction using a structured, Excel-based form. Based on the abstracted data, a list of potential indicators of patients with an anxiety disorder, OCD, or PTSD in need of highly specialized care was generated. No systematic review protocol was registered.

Phase 2. Development of a conceptual framework

In the second phase, a concept mapping study [105] was carried out to inform the generation of scale items. Concept mapping is a mixed-method participatory approach that integrates conventional qualitative group processes (e.g. brainstorming, pile sorting) with multivariate statistical methods of multidimensional scaling in order to depict the composite thinking of participants in a single visual framework ("the concept map"). Experts (clinicians and researchers) in the field of anxiety disorders, OCD, and/or PTSD were invited to participate in the concept mapping process. In total, 147 experts were approached, 34 of which participated in the subsequent stages of vthe concept mapping procedure (i.e. brainstorming and sorting). The concept mapping process and the subsequent data analyses were carried out using Concept Systems software (Concept Systems Incorporated, Ithaca, New York). During the brainstorming stage, participants were asked to review and, when necessary, add additional patient indicators to the list of indicators from the systematic review. Subsequently, participants were asked to sort

the resulting list of indicators into conceptual categories. In order to generate preliminary cluster solutions, the sorting data were analysed using non-metric multidimensional scaling and agglomerative hierarchical cluster analyses. Following procedures recommended by Trochim [112], preliminary cluster solutions were evaluated for their within-cluster coherence of content. In a three-hour consensus meeting, members of the DTAOP consortium were asked to review cluster maps sequentially and select the optimal cluster map solution through an iterative process. The optimal concept map consisted of eight clusters, which, in their turn, consisted of individual potential indicators that jointly contributed to an overarching conceptual domain. For more details on the methods of the second phase see Van Krugten et al. [115; Chapter 3].

Phase 3. Scale generation and evaluation of face validity

Based on the resulting overarching domains (i.e. clusters) of indicators from the concept mapping study, consortium members constructed the draft DTAOP in a three-hour consensus meeting. Since the aim was to develop an easily administrable measure, each of the overarching domains that resulted from the concept mapping study was operationalized into a dichotomous (absent/present) item, resulting in an eight-item draft version of the DTAOP. Response options of the draft DTAOP are "Yes" and "No", scored as 1 and 0, respectively. To ensure face validity, the draft DTAOP was pilot-tested in a sample of 25 outpatients aged 18 years and over with a primary anxiety disorder according to Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV-TR criteria [60]. Clinicians were asked to administer the DTAOP, indicate whether the DTAOP was complete and useful and provide qualitative comments on item clarity. Based on the pilot data, minor changes were made in the wording of some of the items to ensure item clarity.

Note that at the time of data collection, the DSM-IV-TR [60] was in use in The Netherlands, in which obsessive-compulsive disorder (OCD) and post-traumatic stress disorder (PTSD) are classified as anxiety disorders. The DSM-5 [61] chapter on anxiety disorders, however, no longer includes OCD and PTSD. To accommodate future use of the DTAOP under DSM-5 criteria in patients with OCD and PTSD, these diagnoses were separated in the text of the items.

Phase 4. Psychometric evaluation

Study design and population

To evaluate the DTAOP in terms of its psychometric performance, a cross-sectional, multicenter observational study was conducted in nine independent specialized (general psychiatric) and highly specialized psychiatric in-and outpatient clinics in The Netherlands. Between April 2016 and December 2016, a total of 454 adult (aged

18 and older) in-and outpatients with at least one DSM-IV-TR anxiety disorder were evaluated with the DTAOP. Exclusion criteria were: aged younger than 18 years, and no DSM-IV-TR anxiety disorder. The DSM-IV-TR diagnosis was established by administration of the MINI-Plus 5.0.0 [155, 156].

Measures

In addition to the DTAOP, the five-level EuroQol five-dimensional questionnaire (EQ-5D-5L) [157] was administered. The EQ-5D-5L is a five-item generic, preference-based self-report measure to describe and value health-related quality of life (HRQoL). The EQ-5D-5L contains five domains (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and a visual analogue scale (EQ-VAS) for overall health. Each domain is divided into five response options describing the state per domain (no problems, some problems, moderate problems, severe problems, and extreme problems/unable to). An index score can be generated by applying societal preference weights to the health states as completed by the respondent. Based on the Dutch tariff, total scores can range from -0.446 to 1 [158], with higher scores indicating better HrQoL. The EQ-VAS is a vertical scale ranging from zero ("worst imaginable health state") to 100 ("best imaginable health state") on which the respondents are asked to rate their overall health.

Procedures

Following referral from a primary or independent specialized (i.e. secondary) mental health provider, the clinician responsible for intake administered the DTAOP. This was done at the end of the routine intake process, after the diagnostic work-up was completed. The scoring on the DTAOP items, the patients' demographic variables and two feasibility questions were recorded in anonymized, electronic case report forms. Feasibility was operationalized as the clarity of the total set of items (scored with "Yes" or "No") and the time required to complete the DTAOP. Inter-rater reliability was examined in a random sample of 20% of patients using independent, concurrent DTAOP evaluations performed by a set of two permutable clinicians. After consenting to participate, patients were invited to self-complete the EQ-5D-5L to assess the convergent validity. Criterion validity was evaluated in a random subsample of 50% of patients. In the absence of a validated reference test, the clinical judgment was the reference standard for the evaluation of the criterion validity. Based on a review of the patient's medical record, two clinicians independently and blinded to the DTAOP score judged whether the patient needed highly specialized psychiatric treatment. An independent researcher verified the agreement between the clinical judgments, and any discrepancies were resolved by discussion between the two clinicians involved. In a three-hour consensus meeting, consortium members reviewed the results of the

psychometric analyses and made when necessary adjustments to the DTAOP, which resulted in the final version of the DTAOP.

Statistical analysis

Demographic and clinical characteristics and feasibility outcomes (clarity and completion time) were analysed by descriptive statistics. Clarity was considered acceptable if ≥90% of the clinicians evaluated the total set of items of the DTAOP as sufficiently clear. The limit of acceptability of the time required to complete the DTAOP was set at ≤10 minutes. Inter-rater reliability was evaluated by the percent agreement and Krippendorff's alpha [132, 133]. In contrast to the percent agreement, Krippendorff's alpha takes into account the agreement expected by chance and is invariant to the permutation of observers. For each Krippendorff's alpha value a 95% bias corrected confidence interval (CI) was generated by 10,000 bootstrap replications. Although clear rules for determining acceptable reliability are lacking, Krippendorff's alpha values of 0.667 and higher have previously been considered adequate [133]. Following an assessment of data distribution using a Shapiro-Wilk test, Spearman's rank correlation coefficients between total DTAOP scores and EQ-5D-5L index scores and EQ-VAS scores were computed to assess convergent validity. Correlations of 0.10-0.29 were considered weak, 0.30-0.49 moderate and ≥0.50 strong [135]. Since HRQoL was demonstrated to be sensitive to variations in patient factors [159], the DTAOP was hypothesized to have negative correlations with the EQ-5D-5L index and EQ-VAS. In order to evaluate the criterion of the DTAOP, a receiver-operating characteristic (ROC) curve was constructed. To determine the optimal cut-off score for identifying patients with an anxiety disorder, OCD, or PTSD in need of highly specialized care, Youden indices (J = sensitivity$_c$ + specificity$_c$ -1) [136] for a range of cut-off scores were generated. To obtain an optimal trade-off between sensitivity and specificity, the cut-off score that achieved the highest Youden index (i.e. the cut-off score that optimized sensitivity and specificity) was selected as the optimal cut-off score. All analyses were conducted using IBM SPSS (Statistical Package for the Social Sciences) version 20.0 (IBM SPSS Statistics for Macintosh, Armonk, NY: IBM Corp.). Statistical significance was set at $P<0.05$ (two-tailed).

Results

Development

The systematic search identified a total of 4,187 references, of which 34 met the inclusion criteria. Based on the included papers, a list of 46 clinical and socio-demographic indicators of patients with an anxiety disorder, OCD, or PTSD in need for highly special-

ized care was generated. The PRISMA flow chart of the study selection process and the resulting list of indicators are provided in Appendix 5.2 and 5.3, respectively. In the brainstorming stage of the concept mapping procedure, 19 additional potential patient indicators were added to the indicators from the systematic review, resulting in a total of 65 indicators of patients with an anxiety disorder, OCD, or PTSD in need of highly specialized care. The resulting concept map revealed eight overarching domains of indicators: treatment course; socio-demographic and personal factors; psychosocial dysfunctioning; psychosocial factors and compensating individual characteristics; psychiatric comorbidity; severity of the anxiety disorder, OCD, or PTSD, suicidal ideation and self-destructive behaviour; and subtypes of OCD. See Appendix 5.3 for the full list of potential indicators and Appendix 5.4 for the resulting concept map. Based on the concept map, the initial draft of the DTAOP was generated. See Table 5.1 for the abbreviated items of the draft version of the DTAOP. An English translation of the full and final DTAOP is provided in Appendix 5.5.

Table 5.1. The Decision Tool Anxiety Disorders, OCD and PTSD (DTAOP): items, response options and scoring system.

	DTAOP item[a]	**Response options**	**Score**
1	Previous unsuccessful treatment of the current primary diagnosis in specialized care	Yes No	1 0
2	Socio-demographic or personal factors maintaining the anxiety disorder, OCD, or PTSD[b] Example: low IQ, positive family history of anxiety disorders, OCD, or PTSD	Yes No	1 0
3	Treatment-interfering psychosocial dysfunctioning	Yes No	1 0
4	Treatment-interfering psychosocial factors and/or compensating individual characteristics Example: inadequate social support system, poor illness insight, low motivations, low level of perceived self-efficacy	Yes No	1 0
5	Treatment-interfering psychiatric comorbidity	Yes No	1 0
6	Severe anxiety disorder, OCD, or PTSD	Yes No	1 0
7	Acute suicidal ideation and/or self-destructive behaviour	Yes No	1 0
8	≥2 subtypes of OCD	Yes No No OCD	1 0 0

DTAOP = Decision Tool Anxiety Disorders OCD and PTSD; IQ = Intelligence Quotient; Obsessive-Compulsive Disorder (OCD); Post-Traumatic Stress Disorder (PTSD).

[a] Item text is abbreviated. See Appendix 5.5 for an English translation of the full and final DTAOP.

Psychometric evaluation

Patient characteristics

Demographic and clinical characteristics of the total sample and its two subsamples are shown in Table 5.2. The mean (SD) age of the total sample was 35.33 (11.74) years (range=18-83 years), and 67.2% of the sample was female. The mean total DTAOP score was 3.10 (SD=1.80, range=0-7). Mean HRQoL scores as measured by the EQ-5D-5L index and EQ-VAS were 0.50 (SD=0.27) and 58.55 (SD=20.28), respectively. Seven patients were excluded from the data analysis because of missing individual DTAOP items (n=3) or because they were aged younger than 18 years (n=4). The frequency with which the individual items of the DTAOP were present in the total sample are shown in Table 5.3.

Table 5.2. Characteristics of the study sample.

	Total sample	**Inter-rater reliability sample[a]**	**Criterion validity sample[a]**
N	454	87	216
Age, years			
Mean (SD)	35.33 (11.74)	34.61 (10.88)	34.86 (11.11)
Range	18-83	19-60	18-65
Sex (n, %)			
Male	149 (32.8)	29 (33.3)	66 (30.6)
Female	305 (67.2)	58 (66.7)	150 (69.4)
Diagnosis, n (%)			
GAD/phobia	230 (50.7)	49 (56.3)	91 (42.1)
OCD	137 (30.2)	24 (27.6)	72 (33.3)
PTSD	61 (13.4)	7 (8.0)	37 (17.1)
GAD/phobia and OCD	12 (2.6)	3 (3.4)	6 (2.8)
GAD/phobia and PTSD	12 (2.6)	4 (4.6)	9 (4.2)
GAD/phobia, OCD and PTSD	2 (0.4)	-	1 (0.5)
Total DTAOP score			
Mean (SD)	3.10 (1.80)	3.22 (1.74)	3.38 (1.79)
Range	0-7	0-7	0-7
EQ-5D-5L index			
Mean (SD)	0.50 (0.27)[b]	0.52 (0.27)[d]	0.49 (0.28)[e]
Range	-0.30-1.00	-0.11-1.00	-0.30-1.00
EQ-VAS			
Mean (SD)	58.55 (20.28)[c]	62.71 (19.11)[d]	58.77 (21.20)[f]
Range	0-100	20-95	0-100

DTAOP = Decision Tool Anxiety Disorders, OCD and PTSD; EQ-5D-5L = five-level EuroQol five-dimensional questionnaire; EQ-VAS = EuroQol visual analogue scale; GAD = Generalized Anxiety Disorder; OCD = Obsessive-Compulsive Disorder; PTSD = Post-Traumatic Stress Disorder; SD = Standard Deviation.
[a] Part of total sample. [b] N=386. [c] N=371. [d] N=78. [e] N=177. [f] N=165.

Table 5.3. Frequency and percentages with which the items of the DTAOP were present in the total sample (n=454).

	DTAOP item	N	%
1	Previous unsuccessful treatment of the current primary diagnosis in specialized care	189	41.6
2	Socio-demographic or personal factors maintaining the anxiety disorder, OCD, or PTSD	186	41.0
3	Treatment-interfering psychosocial dysfunctioning	229	50.4
4	Treatment-interfering psychosocial factors and/or compensating individual characteristics	203	44.7
5	Treatment-interfering psychiatric comorbidity	214	47.1
6	Severe anxiety disorder, OCD, or PTSD	295	65.0
7	Acute suicidal ideation and/or self-destructive behaviour	26	5.7
8	≥2 subtypes of OCD	65	14.3

DTAOP = Decision Tool Anxiety Disorders, OCD and PTSD; OCD = Obsessive-Compulsive Disorder; PTSD = Post-Traumatic Stress Disorder.

Feasibility

The average time required to complete the DTAOP was 4.6 minutes (i.e. 4 minutes and 37 seconds) (SD=2.62, range=1-20) and the total DTAOP judged as clear in the majority (93.0%) of all evaluations. Nine clinicians expressed concern about the distinctiveness and clarity of item 2 ("Socio-demographic or personal factors maintaining the anxiety disorder, OCD or PTSD") and item 4 ("Treatment-interfering psychosocial factors and/or compensating individual characteristics"). Additionally, eight clinicians suggested the addition of a cut-off score to item 6 ("Severe anxiety disorder, OCD or PTSD") by which the presence or absence of a severe anxiety disorder, OCD or PTSD could be determined. Based on the provided qualitative feedback and further results on the psychometric properties of the DTAOP, consortium members proposed revisions to the wording and instructions of some of the items, resulting in the final version of the DTAOP. See the paragraph "Continued development of the DTAOP" for the description and results of the continued development of the DTAOP.

Reliability

As shown in Table 5.4, the percentage of agreement ranged from 71% to 92%, and Krippendorff's alpha values ranged from 0.4274 (95% CI, 0.2428-0.6015) for item 2 ("Social factors maintaining the anxiety disorder, OCD or PTSD") to 0.8392 (95% CI, 0.7203-0.9401) for item 1 ("Previous unsuccessful treatment of the current primary diagnosis in specialized care"). The Krippendorff's alpha values of items 2 to 6 were below the recommended level of 0.667 [133].

Validity

Consistent with our hypotheses, the DTAOP negatively correlated with the EQ-5D-5L index ($r_s(386)$=-0.413; P<0.001) and EQ-VAS ($r_s(371)$=-0.296; P<0.001). See Figure 5.2

Table 5.4. Inter-rater reliability indices as assessed by percent agreement and Krippendorff's alpha (n=87).

	DTAOP item	% agreement	Krippendorff's alpha (95% CI)
1	Previous unsuccessful treatment of the current primary diagnosis in specialized care	92	0.8392 (0.7203-0.9401)
2	Socio-demographic or personal factors maintaining the anxiety disorder, OCD or PTSD	71	0.4274 (0.2428-0.6015)
3	Treatment-interfering psychosocial dysfunctioning	82	0.6339 (0.4824-0.7810)
4	Treatment-interfering psychosocial factors and/or compensating individual characteristics	80	0.6114 (0.4432-0.7614)
5	Treatment-interfering psychiatric comorbidity	72	0.4346 (0.2417-0.6106)
6	Severe anxiety disorder, OCD or PTSD	83	0.6235 (0.4541-0.7816)
7	Acute suicidal ideation and/or self-destructive behaviour	98	0.7890 (0.4494-1.0000)
8	≥ 2 subtypes of OCD	84	0.8153 (0.7395-0.8865)

CI = Confidence Interval; DTAOP = Decision Tool Anxiety Disorders, OCD and PTSD; OCD = Obsessive-Compulsive Disorder; PTSD = Post-Traumatic Stress Disorder.

and Table 5.5 for the operating characteristics of the DTAOP at various cut-off scores. The area under the curve (AUC) was 0.826 (95% CI, 0.772-0.881; P<0.001) and the Youden index was highest at a cut-off score of ≥4 (J=0.471), with a sensitivity of 0.700 (95% CI, 0.610-0.780) and a specificity of 0.771 (95%, CI 0.674-0.850).

Continued development of the DTAOP

Based on the qualitative feedback (feasibility results) and reliability estimates, consortium members proposed revisions to the wording and instructions of items 2 to 6. To improve the distinctiveness and item clarity of item 2 ("Socio-demographic or personal factors maintaining the anxiety disorder, OCD or PTSD") and 4 ("Treatment-interfering psychosocial factors and/or compensating individual characteristics"), the item wording and instructions of both items were revised. Additionally, cut-off scores by which the presence or absence of treatment-interfering psychosocial dysfunctioning (item 3) and the presence or absence of a severe anxiety disorder, OCD or PTSD (item 6) can be determined were added. Finally, to improve the clarity of item 5 ("Treatment-interfering psychiatric comorbidity"), an item instruction was added by which the presence or absence of a treatment-interfering comorbid psychiatric disorder can be determined. An English translation of the revised and final DTAOP is presented in Appendix 5.5. Although the changes are likely to improve item clarity and subsequently enhance item-level inter-rater reliability, future studies are needed to determine the inter-rater reliability of the newly worded items and instructions.

Figure 5.2. ROC curve for the DTAOP (N=216; AUC=0.826, 95% CI, 0.772-0.881; P<0.001).

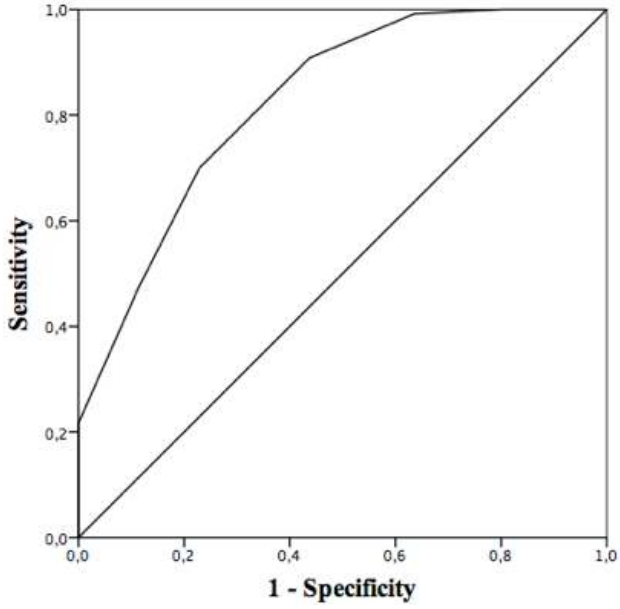


Table 5.5. Operating characteristics of the DTAOP.

DTAOP scale score	Sensitivity (95% CI)	Specificity (95% CI)	Youden index[a]
≥1	1.000 (0.970-1.000)	0.198 (0.124-0.292)	0.198
≥2	0.992 (0.954-1.000)	0.365 (0.269-0.469)	0.357
≥3	0.908 (0.842-0.953)	0.562 (0.457-0.664)	0.470
≥4	0.700 (0.610-0.780)	0.771 (0.674-0.850)	0.471
≥5	0.475 (0.383-0.568)	0.885 (0.804-0.941)	0.360
≥6	0.217 (0.147-0.301)	1.000 (0.962-1.000)	0.217
≥7	0.017 (0.000-0.030)	1.000 (0.962-1.000)	0.017

DTAOP = Decision Tool Anxiety Disorders, OCD and PTSD; CI = Confidence Interval.
[a] Youden index = (sensitivity + specificity) - 1.

Discussion

This study presented the development and psychometric evaluation of the Decision Tool Anxiety Disorders, OCD and PTSD (DTAOP). The DTAOP is an eight-item clinician-administered screening measure designed to facilitate the early identification of patients with an anxiety disorder, OCD or PTSD in need of highly specialized care. Scale items were selected in a sequential mixed-methods approach to allow an in-depth exploration of the factors indicating a need for highly specialized care in patients with an anxiety disorder, OCD or PTSD. In this cross-sectional, multicenter observational study

in patients with a DSM-IV-TR anxiety disorder, the DTAOP demonstrated excellent feasibility and good validity, but weak inter-rater reliability. To improve the item-level inter-rater reliability, revisions and refinements of the wording and instructions were made, resulting in the final version of the DTAOP.

The study was performed in a sample of patients with a DSM-IV-TR anxiety disorder in routine in-and outpatient treatment to maximize external validity and clinical relevance. The clarity of the total set of items was supported by the majority (93.0%) of respondents and the average completion time was 4 minutes and 37 seconds, indicating that the DTAOP is quick to complete. Despite the satisfactory feasibility results, indicating that scoring of the DTAOP was clear and quick on an individual level, the Krippendorff's alpha values of five of the items fell short of the minimum recommended reliability level of 0.667 [133]. Although this may be partly due to the use of a highly rigorous measure for assessing inter-rater reliability [134], the qualitative feedback revealed that revisions and refinements of the wording and instructions of the respective items could improve the inter-rater reliability. Although the changes made to the DTAOP were aimed at improving the item-level inter-rater reliability, future studies are needed to confirm this. In addition, previous work has shown that training and frequent use in daily clinical practice can significantly improve scale reliability [137, 138]. Whether training and frequent use also improves DTAOP inter-rater reliability levels, should, however, be subject to future research as well.

Aggregated DTAOP scores demonstrated meaningful patterns of convergent validity with HRQoL scores as measured by the EQ-5D-5L index and EQ-VAS. DTAOP scores were more strongly associated with EQ-5D-5L index scores than with EQ-VAS scores. The stronger association with EQ-5D-5L index scores could be explained by the fact that both the DTAOP and EQ-5D-5L are scored (DTAOP) or valued (index values EQ-5D-5L) by someone other than the patient, which could have reduced effects of for instance coping and adaptation. The DTAOP also demonstrated good criterion validity (AUC=0.826), indicating that the consensus-based conceptual framework that guided DTAOP development fits the measured construct well.

Main strengths of the present study include the mixed-methods approach used to develop the measure, the large number of examined psychometric properties, and the nationwide representation of participating clinics (nine independent general psychiatric and highly specialized psychiatric in-and outpatient clinics across The Netherlands). Also, to our knowledge, the DTAOP is the first psychometrically validated measure to assess highly specialized care need in patients with an anxiety disorder, OCD, or PTSD. It meets the need for an accurate and easily administrable measure to facilitate the

early identification and referral of patients with an anxiety disorder, OCD, or PTSD in need of highly specialized care. However, several limitations of this study should be noted. First, since the present study represents a first cross-sectional evaluation of the psychometric properties of the DTAOP, future studies are needed to replicate and extend these initial findings. More specifically, important areas for future research include the assessment of the inter-rater reliability of the adapted items, the convergent validity with measures of anxiety disorder, OCD, or PTSD severity and psychosocial functioning, the predictive validity for use in clinical and research settings, and the sensitivity to treatment-related change. In addition, since the present study was not powered to detect differences in psychometric performance of the DTAOP between types of DSM-IV-TR anxiety disorders, future research should assess whether DTAOP performs differently in different types of anxiety disorders. Second, it should be noted that the electronic case report forms in which the scoring of the DTAOP was entered by clinicians, did not allow items to be left unanswered. Hence, an evaluation of the feasibility in terms of missing values could not be performed. Third, in the absence of a standard test for the systematic and standardized early identification of patients with a highly specialized mental healthcare need, the clinical judgement constituted the reference standard for the evaluation of the criterion validity. Although the use of the clinical judgement as the reference standard may have introduced subjective error, effort was made to reduce error by basing the final clinical judgement on dual, independently provided examinations made by two clinicians who were blinded to the index (i.e. DTAOP) score. Fourth, since the aim was to develop an easily administrable measure, the scoring system of the DTAOP was simplified to indicating the 'absence' or 'presence' of the respective clinical (e.g., suicidal ideation) and non-clinical patient factors (e.g. psychosocial factors). However, it should be noted that measuring clinical and non-clinical patient factors is a complex and nuanced matter. Sensitive and valid assessment of the respective factors, and assessment of their possible treatment-interfering effect (items 3-5), may require a more sensitive approach like Likert or even continuous scoring systems. Likewise, the DTAOP was constructed as an unweighted additive scoring system in order to be easily administrable. Although the use of an unweighted additive scoring system enhances the feasibility (i.e. ease of use) of the DTAOP within the context of daily clinical practice, it diminishes the proportional effect of individual items and possible meaningful interactions between items, and may thereby reduce the validity (i.e. precision) of the resultant classification of patients. However, irrespective of the use of a two-point (dichotomous) scoring system and additive score model of unweighted items, the DTAOP demonstrated to be a valid and clinically applicable operationalization of highly specialized care need. Further work could be carried out to establish the effect of different item-level scoring systems and the use of a weighted scoring system on the psychometric properties

of the DTAOP. Fifth, although the established cut-off score of ≥4 is likely to generalize to specialized and highly specialized care settings in The Netherlands due to the national uniform organizational structure and service delivery of psychiatric services, future studies are needed to establish its cross-national robustness. Sixth, although the DTAOP was initially designed for use in specialized mental healthcare centers to inform "step-up" referral decisions to highly specialized mental healthcare centers, the DTAOP might also inform "step-down" referral decisions from highly specialized care back to specialized mental healthcare care. In addition, use of the DTAOP in primary mental healthcare may further enhance the early identification of patients in need of highly specialized care and the timely selection of the optimal initial treatment in these patients. Future studies are required to evaluate the possible added benefit of such broader use of the DTAOP in primary and highly specialized mental health services. Finally, it should be noted that the DTAOP is not designed to replace careful clinical assessment, but is rather intended to provide probable indications of highly specialized care need and should be used as a first step in a more comprehensive assessment. As such, the DTAOP has the potential to aid in the selection of the most appropriate treatment setting for patients on an individual basis, ultimately benefitting the clinical and cost-effectiveness of treatments.

Despite its limitations, this study provides initial support for the psychometric properties of the DTAOP in a sample of patients with a DSM-IV-TR anxiety disorder. The DTAOP demonstrated to be a short and easy scoring, and at a cut-off score of ≥4, valid measure to aid clinicians in the early identification of patients with an anxiety disorder, OCD, or PTSD in need of highly specialized care. Future research is needed to determine the inter-rater reliability of the newly worded items and instructions. Its use in clinical practice will guide in selecting the most appropriate treatment setting, and hence has the potential to benefit treatment outcomes and the efficient use of scarce resources.

Appendix 5.1 – Search strategies

Table 5.1.1. Search strategy Pubmed (NLM).

#	Searches
1	("anxiety disorders"[MeSH Major Topic] OR "panic disorder"[Mesh major topic] OR "agoraphobia"[Mesh major topic] OR "phobic disorders"[mesh major topic] OR "obsessive-compulsive disorder"[MeSH Major Topic] OR "stress disorders, post-traumatic"[MeSH Major Topic] OR "stress disorders, traumatic, acute"[MeSH Major Topic] OR "hyperventilation"[MeSH Major Topic] OR generalized anxiety disorder[tiab] OR social phobia[tiab]) AND (resistant[ti] OR comorbidity[ti] OR comorbid[ti] OR co-morbidity[ti] OR "co morbidity"[ti] OR "co morbid"[ti] OR severe[ti] OR severity[ti] OR complex[ti] OR complexity[ti] OR non response[ti] OR non-response[ti] OR chronic[ti] OR chronicity[ti] OR recurrent[ti] OR recurring[ti] OR recurrence[ti] OR relapse[ti] OR avoidance[ti] OR admission[ti] OR admissions[ti] OR insight[ti] OR duration[ti] OR "failed treatment"[ti] OR "treatment failure"[ti] OR"patient dropouts"[mesh major topic])
2	limit 1 to full text
3	limit 2 to human
4	limit 3 to english or Dutch language
5	limit 101 to yr="2000 -Current"

Table 5.1.2. Search strategy PsycINFO (Ovid).

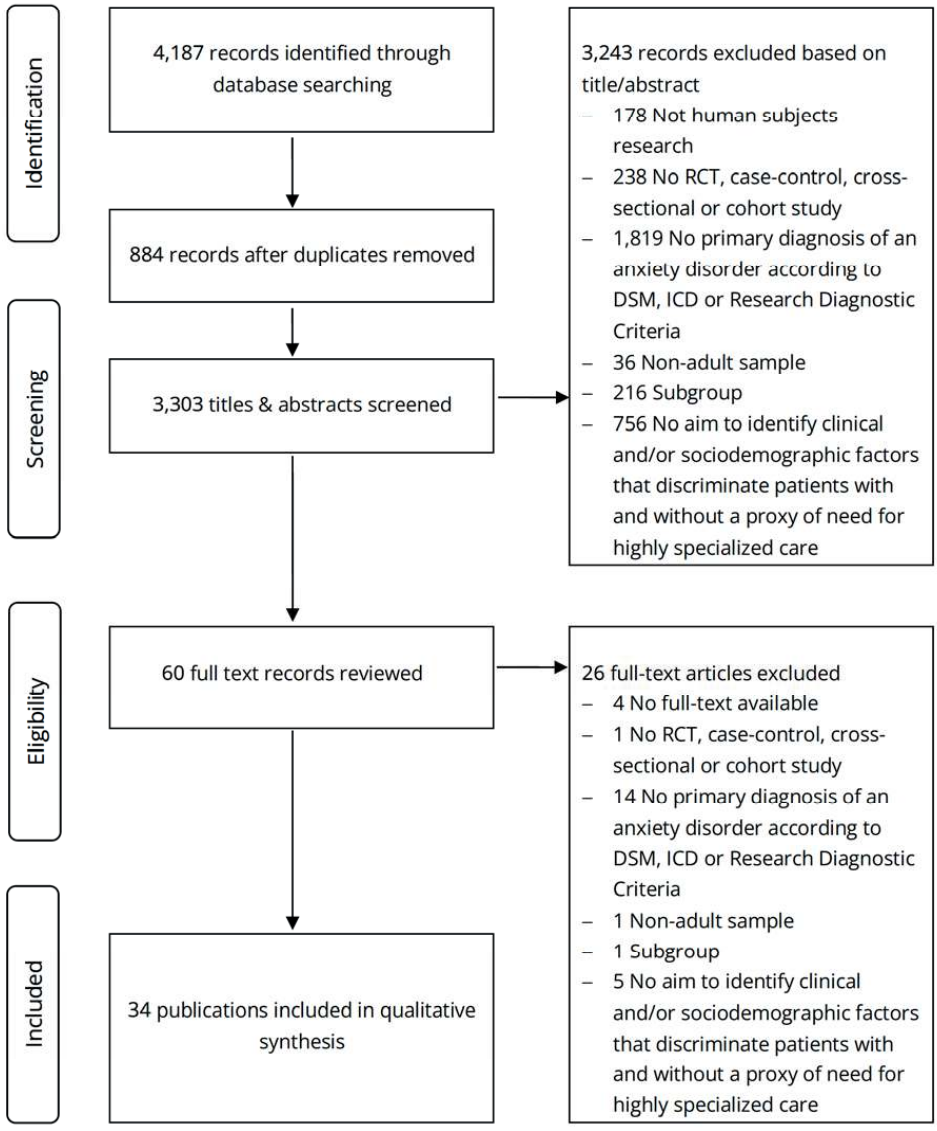
#	Searches
1	("Anxiety disorder" or "disorder, anxiety" or "disorders, Anxiety" or "neuroses, Anxiety" or "Anxiety States, Neurotic" or "Anxiety State, Neurotic" or "Neurotic Anxiety State" or "Neurotic Anxiety States" or "State, Neurotic Anxiety" or "States, Neurotic Anxiety").ab,ti.
2	("Disorder, Panic" or "Disorders, Panic" or "Panic Disorders" or "Panic Attacks" or "Attack, Panic" or "Attacks, Panic" or "Panic Attack").ab,ti.
3	Agoraphobia.ab,ti.
4	(Disorder, Phobic or Disorders, Phobic or Phobic Disorder or Phobic Neuroses or Neuroses, Phobic or Phobias or Phobia or Phobia, School or Phobias, School or School Phobia or School Phobias or Claustrophobia or Claustrophobias or Phobia, Social or Phobias, Social or Social Phobia or Social Phobias).ab,ti.
5	(Disorder, Obsessive-Compulsive or Disorders, Obsessive-Compulsive or Obsessive Compulsive Disorder or Obsessive-Compulsive Disorders or Neurosis, Obsessive-Compulsive or Neuroses, Obsessive-Compulsive or Neurosis, Obsessive Compulsive or Obsessive-Compulsive Neuroses or Obsessive-Compulsive Neurosis or Anankastic Personality or Anankastic Personalities or Personalities, Anankastic or Personality, Anankastic).ab,ti.
6	(Post-Traumatic Stress Disorder or Stress Disorder, Post-Traumatic or Stress Disorders, Post Traumatic or PTSD or Stress Disorder, Post Traumatic or Neuroses, Posttraumatic or Posttraumatic Neuroses or Posttraumatic Stress Disorders or Posttraumatic Stress Disorder or Stress Disorder, Posttraumatic or Stress Disorders, Posttraumatic or Neuroses, Post-Traumatic or Neuroses, Post Traumatic or Post-Traumatic Neuroses or Post-Traumatic Stress Disorders or Post Traumatic Stress Disorders or Chronic Post-Traumatic Stress Disorder or Chronic Post Traumatic Stress Disorder or Delayed Onset Post-Traumatic Stress Disorder or Delayed Onset Post Traumatic Stress Disorder or Acute Post-Traumatic Stress Disorder or Acute Post Traumatic Stress Disorder).ab,ti.
7	(Stress Disorders, Acute or Acute Stress Disorder or Stress Disorder, Acute or Acute Stress Disorders).ab,ti.
8	Hyperventilation.ab,ti.
9	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
10	resistant.ti.

Table 5.1.2. Search strategy PsycINFO (Ovid). (continued)

#	Searches
11	Comorbidity.ti.
12	Comorbid.ti.
13	Co-morbidity.ti.
14	Co morbidity.ti.
15	co morbid.ti.
16	Severe.ti.
17	Severity.ti.
18	Complex.ti.
19	Complexity.ti.
20	Non response.ti.
21	Non-response.ti.
22	Chronic.ti.
23	Chronicity.ti.
24	Recurrent.ti.
25	Recurring.ti.
26	Recurrence.ti.
27	Relapse.ti.
28	Avoidance.ti.
29	Admission.ti.
30	Admissions.ti.
31	Insight.ti.
32	Duration.ti.
33	failed treatment.ti.
34	treatment failure.ti.
35	(Dropout, Patient or Dropouts, Patient or Patient Dropout or Dropout Characteristics or Characteristic, Dropout or Characteristics, Dropout or Dropout Characteristic or Dropouts or Dropout).ab,ti.
36	10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35
37	9 and 36
38	Limit 37 to full text
39	Limit 38 to human
40	Limit 39 to dutch or english language
41	Limit 40 to yr="2000 -Current")

Appendix 5.2 – PRISMA flow chart

Figure 5.2.1. Flow chart of study selection process.



RCT, Randomized controlled trial; DSM, Diagnostic and Statistical Manual of Mental Disorders; ICD, International Classification of Diseases.

Appendix 5.3 – Concept map clusters and indicators

Table 5.3.1. Concept map clusters and indicators.

Cluster	Indicator	Systematic Review[a]	Brainstorm[b]
1. Treatment course			
1.	>1 time relapse		X
15.	Repeated treatments without remission		X
17.	Treatment resistant		X
33.	Partial remission	X	
35.	Earlier onset age	X	
45.	No current pharmacological treatment	X	
51.	Chronic course	X	
56.	Higher level of pretreatment symptoms	X	
2. Socio-demographic and personal factors			
11.	Lower intellectual functioning		X
14.	Low level of motivation		X
24.	Younger age	X	
34.	Having no partner	X	
37.	Female	X	
47.	Higher self-transcendence score	X	
48.	Fewer years of education	X	
52.	OCD in first-degree relatives	X	
57.	Perceived criticism on patient from family members	X	
60.	Perceived criticism	X	
3. Psychosocial dysfunctioning			
2.	Worse functioning		X
5.	Extensive consequential damages		X
12.	Severe stagnation in multiple life domains		X
25.	Disabilities in physical functioning		X
31.	Level of functioning	X	
43.	Unemployment	X	
63.	Inability to work due to illness	X	
4. Psycho-social factors and compensating individual			
4.	Low self-efficacy		X
10.	High tendency to avoid anxiety		X
13.	Lack of compensating competencies		X
29.	Stressful life events	X	
55.	Psychosocial difficulties	X	
59.	Poor insight	X	
64.	Less vitality	X	

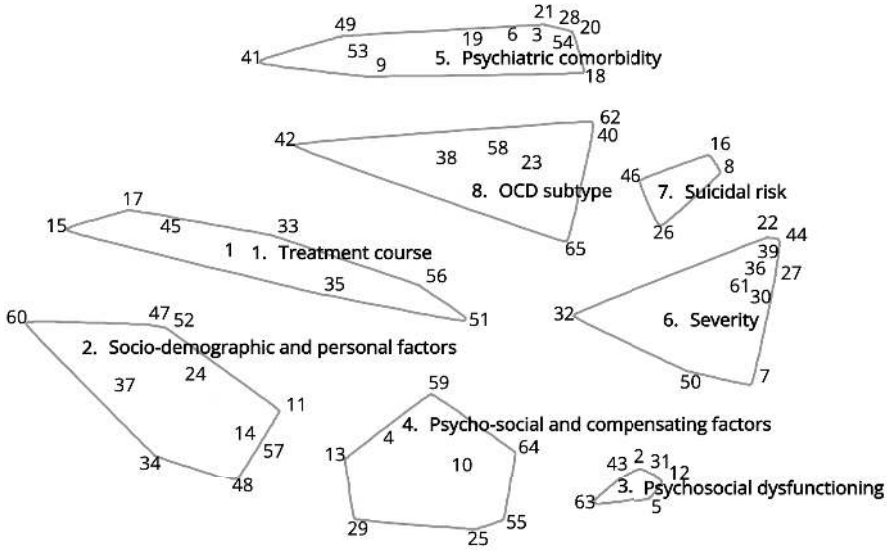
Table 5.3.1. Concept map clusters and indicators. (continued)

Cluster	Indicator	Systematic Review[a]	Brainstorm[b]
5. Psychiatric comorbidity			
3.	Multiple comorbid diagnoses		X
6.	OCD with comorbid TIC-disorders		X
9.	Less common anxiety complaints such as conversion		X
18.	Severe eating disorder		X
19.	Bodyd dismorphic disorder		
20.	Comorbid severe depression	X	
21.	Comorbid personality disorder	X	
28.	Comorbid anxiety disorder	X	
41.	Avoidant personality disorder	X	
49.	Bipolar disorder	X	
53.	Obsessive-compulsive personality disorder	X	
54.	Comorbid alcohol or other substance use disorder	X	
7. Severity			
7.	High level of distress		X
22.	Severe OCD	X	
27.	Severe anxiety symptoms	X	
30.	Severity of complaints	X	
32.	Level of neuroticism	X	
36.	Severity of avoidance in PTSD	X	
39.	Severity of PTSD	X	
44.	Severity of hoarding	X	
50.	Higher level of arousal	X	
61.	Higher severity of anxiety	X	
8. Suicidal risk			
8.	Presence of suicidal risks		X
16.	Severe self-destructive behaviour		X
26.	Severity of childhood trauma	X	
46.	History of self-harm	X	
9. OCD subtypes			
23.	Contamination fears and washing compulsions	X	
38.	Aggressive obsessions	X	
40.	Higher number of OCD-subtypes	X	
42.	Forbidden thoughts (sexual, religious and aggressive obsessions)	X	
58.	Somatic obsessions	X	
62.	Severe depression	X	
65.	Higher level of physical aggression	X	
Total		**46**	**19**

[a] Indicators identified in the systematic review (phase 1).
[b] Indicators generated in the brainstorming stage of the concept mapping procedure (phase 2).

Appendix 5.4 – Concept map

Figure 5.4.1. Concept map of the eight overarching domains of patients with an anxiety disorder in need of highly specialized care (stress value = 0.298).



The numbers on the concept map correspond to the indicators that were sorted into each category (see Appendix 5.3 for an overview of the indicators). Indicators that are closer together indicate higher degrees of similarity based on sorting.

Appendix 5.5 – Decision Tool Anxiety Disorders, OCD and PTSD

Figure 5.5.1A. English translation of the full and final DTAOP (part 1 of 2).

Decision Tool Anxiety Disorders (DTAD)

Name of patient:

Date:

Name of clinician:

		Comment
1. Have there been any unsuccessful treatments in specialized mental healthcare for the primary diagnoses?	○ yes ○ no	
2. Are there any social factors contributing to the anxiety disorder that are hard to influence? Note: Also think of low education, unemployment, little or no support system, and a dysfunctioning family system.	○ yes ○ no	
3. Does the patient exhibit severe psychosocial dysfunctioning that interferes with the anxiety, OCD, or PTSD treatment? • Note: • GAF≤50 or WHODAS≥130 is an indication of severe dysfunctioning. • There is interference if the degree of psychosocial dysfunctioning complicates the clinical presentation of the primary diagnosis, or the treatment of the primary diagnosis.	○ yes ○ no	
4. Does the patient have a disadaptive coping style that interferes with the anxiety, OCD, or PTSD treatment? Note: Think of low motivation, lack of compensating individual characteristics, and a low level of perceived self-efficacy.	○ yes ○ no	
5. Does the patient have at least one diagnosed comorbid psychiatric disorder that interferes with the anxiety, OCD, or PTSD treatment? Note: • Also think of personality disorders, development disorders, addiction, and intellectual disabilities. • There is interference if the diagnosed comorbid disorder complicates the clinical presentation of the primary diagnosis, or the treatment of the primary diagnosis.	○ yes ○ no	
6. Does the patient have a severe anxiety disorder, OCD, or PTSD? Note: besides a clinical impression of severity does the patient score high on one of the following measures: • General measures: SCL-90 high or very high in comparison to a normative sample of outpatients; BSI high or very high in comparison to a normative sample of outpatients. • Anxiety disorder: BAI ≥26. • OCD: Y-BOCS ≥24; diminished/no sense of reality. • PTSD: CAPS-5 average item score >3; PCL-5 average item score >3; DSM-5 severe.	○ yes ○ no	
7. Does the patient have acute suicidal ideation and/or self-destructive behaviour?	○ yes ○ no	

P.T.O.

Figure 5.5.1B. English translation of the full and final DTAOP (part 2 of 2).

Decision Tool Anxiety Disorders (DTAD)

8. In case of OCD, are there 2 or more subtypes present?
   Examples of OCD subtypes are:
   - Compulsive washing
   - Compulsive checking
   - Compulsive hoarding
   - Obsession with symmetry, ordering/arranging, or counting
   - Aggressive, religious, or sexual intrusion

   ○ yes
   ○ no
   ○ no OCD

   Comment

Total amount of positive (=yes) scores ≥ 4?

Yes → indicated for highly specialized care on the basis of the DTAD
No → not indicated for highly specialized care on the basis of the DTAD

6

Development and psychometric evaluation of the Transdiagnostic Decision Tool: Matched care for patients with a mental disorder in need of highly specialized care

van Krugten FCW
van der Feltz-Cornelis CM
Boeschoten MA
van Broeckhuysen-Kloth SAM
van Eck van der Sluijs JF
van Ee E
van Es SM
Schoorl M
Tak LM
Brouwer WBF
Hakkaart-van Roijen L

BJPsych Open. 2020:6(5).

Abstract

Objectives

Early identification of patients with mental health problems in need of highly specialized care could enhance the timely provision of appropriate care and improve the clinical and cost-effectiveness of treatment strategies. Recent research on the development and psychometric evaluation of diagnosis-specific decision support algorithms suggested that the treatment allocation of patients to highly specialized mental healthcare settings may be guided by a core set of transdiagnostic patient factors. The aim of this study, therefore, was develop and psychometrically evaluate a Transdiagnostic Decision Tool to facilitate the uniform assessment of highly specialized mental healthcare need in heterogeneous patient groups.

Methods

The Transdiagnostic Decision Tool was developed based on an analysis of transdiagnostic items of earlier developed diagnosis-specific Decision Tools. The Transdiagnostic Decision Tool was psychometrically evaluated among 505 patients with a somatic symptom disorder or post-traumatic stress disorder. Feasibility, inter-rater-reliability, convergent validity and criterion validity were assessed. In order to evaluate convergent validity, the five-level EuroQol five-dimensional questionnaire (EQ-5D-5L) and the ICEpop CAPability measure for Adults (ICECAP-A) were administered.

Results

The six-item clinician-administered Transdiagnostic Decision Tool demonstrated excellent feasibility and acceptable inter-rater reliability. Spearman's rank correlations between the Transdiagnostic Decision Tool and ICECAP-A (-0.335), EQ-5D-5L index (-0.386) and EQ-5D-VAS (-0.348) supported convergent validity. The area under the curve was 0.81 and a cut-off value of ≥3 was found to represent the optimal cut-off value.

Conclusions

The Transdiagnostic Decision Tool demonstrated solid psychometric properties and showed promise as a measure for the early detection of patients in need of highly specialized mental healthcare.

Introduction

Although the efficacy of psychological interventions for the treatment of a wide range of mental health problems is well established [160, 161], a significant number of patients require multiple treatment steps to achieve an adequate treatment response [120]. An inadequate response to initial treatment, in turn, is associated with higher relapse rates, chronicity [120], and substantial societal costs [163]. Against this background, and given the increasing prevalence [1] and high associated costs [164] of mental health problems, the importance of matching patients to the most appropriate level and type of initial care is increasingly recognized [e.g., 165].

The matched care approach, in which pre-treatment patient characteristics are used to match patients to the level of care that is likely to be most beneficial to them [28], has the potential to improve the clinical effectiveness and cost-effectiveness of treatment strategies [166]. Matched care has been demonstrated to be an appropriate and effective approach in patients with mental health problems attending the primary care setting [167, 168], the occupational healthcare setting [169], and the outpatient general hospital setting [170], but is likely to be most beneficial for the subgroup of patients in need of highly specialized mental healthcare. Often, these patients demonstrate low response and high relapse rates after initial treatment [171, 172], and require additional treatment steps as the result. The provision of matched care in this subgroup is therefore warranted, but strongly relies on the ability to identify these patients and therefore the availability of pre-treatment assessment tools and decision guidelines to accurately match the initial treatment to the individual patient needs [122, 124].

Recent initiatives to inform treatment decisions by pre-treatment patient characteristics include the development of diagnosis-specific Decision Tools for the diagnostic groups personality disorders [173], eating disorders [174], unipolar depression [128; Chapter 4] and anxiety disorders, OCD and PTSD [175; Chapter 5]. Decision Tools are brief, clinician-administered instruments, especially designed to identify patients in need of highly specialized care during the diagnostic phase. Decision Tools items include pre-treatment patient characteristics such as the absence or presence of psychiatric or somatic comorbidity, and the total score is an indicator of the need for highly specialized care. The diagnosis-specific Decision Tools demonstrated solid psychometric properties [128, 173-175; Chapters 4 and 5], and are used in psychiatric specialized centers to enhance the early identification of patients with a highly specialized mental healthcare need. The development and psychometric evaluation of the diagnosis-specific Decision Tools suggested that the allocation of patients to highly

specialized mental healthcare settings may be guided by a core set of transdiagnostic patient factors. Building on the theoretical foundations of and insights from the development of these diagnosis-specific Decision Tools, the aim of this study was to explore the possibility of developing a transdiagnostic Decision Tool for use in heterogeneous patient groups, in patients with a diagnosis for which no diagnosis-specific Decision Tool is available, and in patients without a clear primary diagnosis. Such a tool could enhance the systematic and standardized early identification of patients with a highly specialized mental healthcare need, which, in turn may enhance treatment outcomes in patients with severe and complex mental health problems.

Methods

Definition of terms and Transdiagnostic Decision Tool development

Highly specialized mental healthcare (i.e. tertiary mental healthcare) is the care provided by highly trained professionals to individuals with mental health problems that are complex and refractory to interventions provided in specialized (e.g. secondary) mental healthcare settings such as community mental health centers and general hospitals [24, 25]. Given the level of necessary staff expertise, assessment, and resources, highly specialized mental healthcare is often, but not per definition, provided in mental healthcare centers affiliated with academic medical settings [25].

In order to enhance the early identification and adequate management of patients with mental health problems in need of highly specialized care, the following four diagnosis-specific decision support algorithms were developed: the Decision Tool Personality Disorders (DTPD) [173], the Decision Tool Eating Disorders (DTED) [174], the Decision Tool Unipolar Depression (DTUD) [128; Chapter 4], and the Decision Tool Anxiety Disorders, OCD and PTSD (DTAOP) [175; Chapter 5]. Building on the theoretical foundations of and insights from the development and psychometric evaluation of these diagnosis-specific Decision Tools, the Transdiagnostic Decision Tool was developed for use in heterogeneous patient groups, in patients with a diagnosis for which no diagnosis-specific Decision Tool is available, and in patients without a clear primary diagnosis. The tool was initially intended for use in the diagnostic phase in specialized mental healthcare centers in order to optimize the clinical decision-making process in the referral of patients with mental health problems to highly specialized care. Its use does not have to be restricted to this setting, however. The Transdiagnostic Decision Tool was developed by the Transdiagnostic Decision Tool Consortium, comprising sixteen leading mental health experts (psychiatrists and psychologists), two academics, and two patient representatives.

The development process of the Transdiagnostic Decision Tool consisted of three consecutive phases. In the first phase, the overlapping patient criteria in the diagnosis-specific Decision Tools were established. In the second phase, consortium members generated the draft Transdiagnostic Decision Tool through operationalization of each of the criteria identified in the first phase. In the third phase, a pilot study was carried out in 34 patients with a Diagnostic and Statistical Manual of Mental Disorders (DSM)-5 [61] diagnosis of post-traumatic stress disorder (PTSD) (N=10), somatic symptom disorder (SSD) (N=10), unipolar depression (N=5), anxiety disorder (N=2), eating disorder (N=3), personality disorder (N=3), or psychotic disorder (N=1) who were referred for treatment to either a specialised or highly specialised treatment centre in the Netherlands. Clinicians were asked to complete the draft version of the Transdiagnostic Decision Tool and answer ques- tions regarding its feasibility. Feasibility questions included the total time required to complete the tool and the clarity of the item wording and the tool in total.

Evaluation of psychometric properties

Study design and population

In order to evaluate the psychometric properties of the Transdiagnostic Decision tool, a cross-sectional, observational multicenter study was carried out in eight specialized (general psychiatric) and highly specialized (i.e. tertiary) mental healthcare clinics in The Netherlands under routine care conditions. To facilitate the comparison of psychometric properties between diagnoses groups and evaluate the transdiagnostic robustness of the Transdiagnostic Decision tool, the study was carried out in two distinct diagnoses groups. The study population consisted of 505 adult (18 years and older) psychiatric outpatients with either a primary diagnosis of SSD or a primary diagnosis of PTSD according to DSM-5 [61] criteria. Written informed consent was obtained from all patients. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. All procedures involving human subjects/patients were approved by the Medical Ethical Committee of the Erasmus University Medical Centre Rotterdam, The Netherlands (MEC-2017-051).

Measures

In addition to the Transdiagnostic Decision Tool that was completed by the clinician, participants also completed a number of self-report instruments.

– The *five-level EuroQol five-dimensional questionnaire* (EQ-5D-5L) [157] is a generic, standardized, self-administered measure of health-related quality of life (HRQoL). The EQ-5D-5L comprises two parts: a descriptive system and a visual analogue

scale (EQ-VAS). The descriptive system consists of five items, covering five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), each with five response levels (no problems, some problems, moderate problems, severe problems, and extreme problems/unable to). The answers on the descriptive system can be converted into a single preference-based summary index score (the EQ-5D-5L index) by applying societal preference weights to the self-classified health states. Based on the Dutch national value set, EQ-5D-5L index scores can range from -0.446 (representing the worst health state) to 1 (representing the best health state) [158]. The second part of the EQ-5D-5L, the EQ-VAS, records the respondent's current self-rated health on a 20 cm vertical scale ranging from zero ("the worst health you can imagine") to 100 ("the best health you can imagine").

– The *ICEpop CAPability measure for Adults* (ICECAP-A) [162] is a generic, standardized, self-administered measure of capability wellbeing for use in the adult population. The descriptive system consists of five items, covering five dimensions (stability, attachment, autonomy, achievement, and enjoyment), each with four response levels. Responses to the descriptive system can be converted into a single summary index by applying societal preference weights to the self-classified capability states. The ICECAP-A index can range from 0 (representing the absence of capability) to 1 (representing full capability) [176].

Procedures

From March 2017 through March 2018, patients were enrolled in the study at eight specialized (general psychiatric) and highly specialized (i.e. tertiary) mental healthcare clinics in The Netherlands. During the intake interview, clinicians rated each participating patient on the Transdiagnostic Decision Tool and entered the scoring on the Transdiagnostic Decision Tool, as well as demographic and clinical characteristics (sex, age, country of origin, primary diagnosis) and two questions regarding the feasibility of the Transdiagnostic Decision Tool into web-based case report forms. Feasibility was operationalized as the total administration time of the Transdiagnostic Decision Tool, the clarity of the total set of items (scored as "Yes" or "No") and the percentage of missing values. In order to evaluate the inter-rater reliability, a random subsample of 28% of patients was rated on the Transdiagnostic Decision Tool by a second clinician present at the intake interview. During the intake interview, patients completed a three-page questionnaire, including the EQ-5D-5L and the ICECAP-A to assess the convergent validity. Based on the patients' preference, the EQ-5D-5L was provided in Dutch, English, French, or Arabic and the ICECAP-A in Dutch or English. Criterion validity was evaluated in a random subsample of 59% of patients by comparing the total Transdiagnostic Decision Tool score with the clinical judgment of senior clinicians. Two clinicians independently and blinded to the individual scores on the Transdiagnostic

Decision Tool rated whether the patient was in need of highly specialized care (scored with "Yes" or "No"). An independent researcher verified the consistency between the judgments, and disagreements were resolved by discussion or through third party consultation.

Statistical analysis

Demographic and clinical characteristics of the study sample and feasibility data were analysed using descriptive statistics. In line with previous Decision Tool research [128, Chapter 4; 175, Chapter 5], criteria for feasibility success were set at a mean administration time of ≤10 minutes, content clarity judged as 'clear' in ≥90% of all evaluations, and ≤5% of missing item responses. To assess the inter-rater reliability, Krippendorff's alpha reliability coefficients [132, 133] were calculated for each of the individual items, and the total Transdiagnostic Decision Tool score. The minimum acceptable reliability level was set at 0.667 [133]. Following Shapiro-Wilk tests of normality, non-parametric Spearman's rank correlations between the total Transdiagnostic Decision Tool scores and EQ-5D-5L index, EQ-5D-5L VAS, and ICECAP-A scores were computed to assess convergent validity. Correlations of 0.10-0.29, 0.30-0.49 and ≥0.50 were considered weak, moderate, and strong, respectively [135]. Transdiagnostic Decision Tool scores were expected to have a moderate negative correlation with HRQoL (EQ-5D-5L) and wellbeing (ICECAP-A) scores. Receiver-operating characteristic (ROC) curves were generated to assess the criterion validity and to determine the optimal cut-off score. Areas under the ROC curves (AUCs) were generated to summarize the discriminative accuracy of the Transdiagnostic Decision Tool. In order to determine the optimal cut-off score, a Youden index ($J = (\text{sensitivity}_c + \text{specificity}_c) - 1$) [136] was calculated for each possible cut-off score. The cut-off score that corresponded to the highest Youden index was selected as the optimal cut-off score. All statistical analyses were carried out both for the total sample and for each diagnostic group, and conducted using IBM SPSS (Statistical Package for the Social Sciences) version 24.0 (SPSS Inc., IBM Corporation, Armonk, New York, USA). Significance levels were set at $P<0.05$ (two-tailed).

Results

Scale development and preliminary evaluation of the criterion validity

Analysis of the overlapping criteria of the diagnosis-specific Decision Tools revealed the following five transdiagnostic criteria to detect patients with a highly specialized care need: high severity level of the primary diagnosis, treatment-interfering psychiatric comorbidity, treatment-interfering somatic comorbidity, treatment-interfering psychosocial dysfunctioning, and previous unsuccessful treatment of the current primary

diagnosis in specialized mental healthcare (see Appendix 6.1 for the primary items of the diagnosis-specific Decision Tools). In a consensus meeting, consortium members added the criterion "Severe or longstanding childhood trauma" to the initial list of five criteria given the prognostic importance of this criterion in patients with mental health problems. In line with the diagnosis-specific Decision Tools [128, 173-175; Chapters 4 and 5], each of the transdiagnostic criteria was operationalized into a dichotomous (item present or not) scale item, resulting in a six-item draft version of the Transdiagnostic Decision Tool. Based on the data of the pilot study, no adjustments to the wording of the items were needed. The items, response options, and scoring system of the Transdiagnostic Decision tool are presented in Table 6.1. An English translation of the complete Transdiagnostic Decision Tool is presented in Appendix 6.2.

Table 6.1. Items, response options and scoring system of the Transdiagnostic Decision Tool.

	Item[a]	Response options	Score
1	Severe primary diagnosis	Yes No	1 0
2	Treatment-interfering psychiatric comorbidity	Yes No	1 0
3	Treatment-interfering somatic comorbidity	Yes No	1 0
4	Treatment-interfering psychosocial dysfunctioning	Yes No	1 0
5	Severe or longstanding childhood trauma	Yes No	1 0
6	Previous unsuccessful treatment of the current primary diagnosis in specialized care	Yes No	1 0

[a] Item text is abbreviated. An English translation of the complete Transdiagnostic Decision Tool is presented in Appendix 6.2.

Psychometric evaluation

In total, 505 patients were enrolled in the study. The demographic and clinical characteristics of the study population are presented in Table 6.2. The mean age of the patients was 41.20 years (SD=12.44; range=18-79), 281 patients (55.6%) were female, and the majority of patients (71.1%) were of Dutch origin. At presentation, 234 (46.3%) patients had a primary diagnosis of SSD, and 271 (53.7%) had a primary diagnosis of PTSD. The mean total Transdiagnostic Decision Tool score was 2.52 (SD=1.76; range=0-6). Mean self-reported HRQoL and wellbeing scores as measured by the EQ-5D-5L, EQ-5D-VAS, and ICECAP-A were 0.40 (SD=0.30; range=-0.35-1.00), 49.68 (SD=19.74; range=0.0-100.0), and 0.58 (SD=0.20; range=0.00-0.97), respectively.

Table 6.2. Demographic and clinical characteristics of the study sample.

	Total sample	IRR sample[a]	Criterion validity sample[a]
N	505	140	298
Age, years			
Mean (SD)	41.20 (12.44)	41.94 (13.21)	41.37 (12.47)
Range	18-79	18-79	18-79
Sex (n, %)			
Male	224 (44.4)	60 (42.9)	116 (38.9)
Female	281 (55.6)	80 (57.1)	182 (61.1)
Country of origin (n, %)			
The Netherlands	359 (71.1)	109 (77.9)	201 (67.4)
Surinam	19 (3.8)	3 (2.1)	10 (3.4)
Turkey	16 (3.2)	2 (1.4)	13 (4.4)
Morocco	14 (2.8)	4 (2.9)	11 (3.7)
Iraq	12 (2.4)	3 (2.1)	10 (3.4)
Syria	11 (2.2)	2 (1.4)	7 (2.3)
Afghanistan	8 (1.6)	2 (1.4)	2 (0.7)
Other	65 (12.9)	15 (10.7)	43 (14.4)
Missing	1 (0.2)	0 (0.0)	1 (0.3)
Diagnosis (n, %)			
SSD	234 (46.3)	87 (62.1)	155 (52.0)
PTSD	271 (53.7)	53 (37.9)	143 (48.0)
Total Decision Tool score			
Mean (SD)	2.52 (1.76)	2.66 (1.83)	2.62 (1.70)
Range	0-6	0-6	0-6
EQ-5D-5L index			
Mean (SD)	0.40 (0.30)	0.40 (0.31)	0.37 (0.31)
Range	-0.35-1.00	-0.35-1.00	-0.35-1.00
Missing (n, %)	20 (4.0)	3 (2.1)	10 (3.4)
EQ-VAS			
Mean (SD)	49.68 (19.74)	48.45 (19.07)	47.11 (19.11)
Range	0.0-100.0	0.00-100.00	0.00-90.00
Missing (n, %)	20 (4.0)	3 (2.1)	10 (3.4)
ICECAP-A index			
Mean (SD)	0.58 (0.20)	0.61 (0.20)	0.58 (0.21)
Range	0.00-0.97	0.08-0.97	0.00-0.97
Missing (n, %)	23 (4.6)	3 (2.1)	12 (4.0)

EQ-5D-5L = Five-level EuroQol five-dimensional questionnaire; ICECAP-A = ICEpop CAPability measure for Adults; IRR = Inter-Rater Reliability; PTSD = Post-Traumatic Stress Disorder; SD = Standard Deviation; SSD = Somatic Symptom Disorder.

[a] Part of total sample.

Feasibility

Mean administration time of the Transdiagnostic Decision Tool was 6.9 minutes (SD=4.2; range=1-30), and the total set items was evaluated as 'clear' in a vast majority of the evaluations (96.6%). The mean administration time was significantly lower (P<0.001) in patients with SSD (5.6 min; SD=3.1) than in patients with PTSD (8.0 min; SD=4.6). The percentage of missing item responses ranged from 0.0% (item 6) to 1.5% (item 5) (mean=0.8%).

Inter-rater reliability

As shown in Table 6.3, Krippendorff's alpha values ranged from 0.724 (95% CI=0.581-0.841) for item 4 ("Psychosocial dysfunctioning") to 0.848 (95% CI=0.731-0.938) for item 5 ("Childhood trauma") in the total inter-rater reliability sample. In the SSD subsample, the Krippendorff's alpha values of item 3 ("Somatic comorbidity") and item 4 ("Psychosocial dysfunctioning") fell short of the recommended reliability level of 0.667 [133]. All other Krippendorff's alpha values of the individual items and the total Transdiagnostic Decision Tool score exceeded the recommended reliability level.

Table 6.3. Krippendorff's alpha values of the Transdiagnostic Decision Tool.

Item		Total IRR sample (N=140)	SSD (N=87)	PTSD (N=53)
1	Severity	0.733 (0.582-0.868)	0.748 (0.614-0.871)	0.704 (0.552-0.843)
2	Psychiatric comorbidity	0.754 (0.618-0.879)	0.720 (0.568-0.849)	0.763 (0.630-0.877)
3	Somatic comorbidity	0.753 (0.611-0.886)	0.655 (0.498-0.791)[a]	0.941 (0.846-1.000)
4	Psychosocial dysfunctioning	0.724 (0.581-0.841)	0.614 (0.446-0.774)[a]	0.870 (0.761-0.957)
5	Childhood trauma	0.848 (0.731-0.938)	0.871 (0.765-0.957)	0.805 (0.681-0.900)
6	Previous treatment	0.757 (0.614-0.886)	0.700 (0.537-0.838)	0.833 (0.713-0.934)
Total Decision Tool score		0.771 (0.724-0.815)	0.732 (0.677-0.784)	0.808 (0.759-0.853)

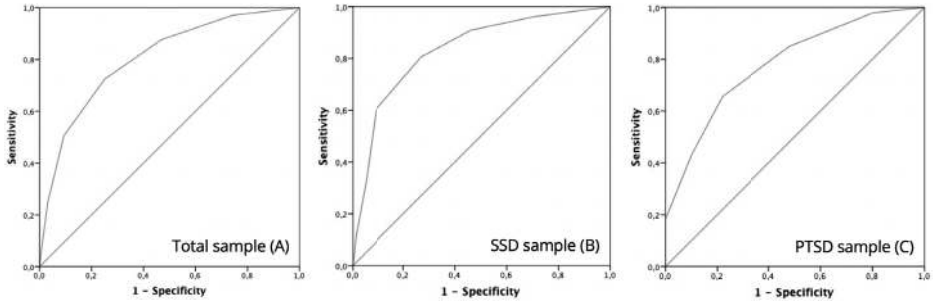
IRR = Inter-Rater Reliability; PTSD = Post-Traumatic Stress Disorder; SSD = Somatic Symptom Disorder.
[a] Below the recommended level of 0.667 [133].

Validity

As hypothesised, Transdiagnostic Decision Tool sum scores negatively correlated with HRQoL and wellbeing scores as measured by the EQ-5D-5L ($r_s(485)=-0.386$; P<0.001), EQ-5D-VAS ($r_s(485)=-0.348$; P<0.001), and ICECAP-A ($r_s(485)=-0.335$; P<0.001). As shown in Figure 6.1, the AUC in the total criterion validity sample (N=298) was 0.81 (95% CI=0.76-0.86; p<0.001). The AUC in the SSD and PTSD subsamples were 0.84 (95% CI=0.77-0.90; p<0.001), and 0.78 (95% CI=0.71-0.86; p<0.001), respectively. The accuracy indices for various cut-off values of the Transdiagnostic Decision Tool are presented in Table 6.4. Based on the highest Youden index (J_{max}) of 0.474 (sensitivity 72.4%; specificity 75.0%), the optimal cut-off value for the Transdiagnostic Decision

Tool was ≥3 in the total criterion validity sample (N=298). The optimal cut-off value of ≥3 was also found in the SSD (J_{max}=0.536) and PTSD (J_{max}=0.436) subsample.

Figure 6.1. ROC curves for the Transdiagnostic Decision Tool.



(A) Total criterion validity sample (AUC=0.81, 95% CI 0.76-0.86; p<0.001) (N=298)
(B) SSD subsample (AUC=0.84, 95% CI 0.77-0.90; p<0.001) (N=155)
(C) PTSD subsample (AUC=0.78, 95% CI 0.71-0.86; p<0.001) (N=143)

Table 6.4. Accuracy indices of the Transdiagnostic Decision Tool in the total criterion validity sample (N=298).

DT Scale Score	Sensitivity		Specificity		Youden index (J)[a]
	%	95% CI	%	95% CI	
≥1	97.1	93.3-99.0	25.8	18.5-34.3	0.228
≥2	87.6	81.7-92.2	53.1	44.1-62.0	0.408
≥3	72.4	65.0-78.9	75.0	66.6-82.2	0.474
≥4	50.6	42.8-58.3	90.6	84.2-95.1	0.412
≥5	24.7	18.4-31.9	96.9	92.2-99.1	0.216
6	7.1	3.7-12.0	99.2	95.7-100.0	0.068

CI = Confidence Interval; DT = Decision Tool.
[a] Youden index = (sensitivity + specificity) - 1.

Discussion

This paper reports on the development and psychometric evaluation of a measure aimed to aid clinicians in the early identification of patients with mental health problems in need of highly specialized care, the Transdiagnostic Decision Tool. Items of the Transdiagnostic Decision Tool were established through identification of overlapping criteria in previously developed diagnosis-specific Decision Tools. Overall, the results of the present study suggest that the Transdiagnostic Decision Tool is a psychometrically sound, and with the establishment of a cut-off score, promising tool for the early identification of patients with mental health problems in need of highly specialized care.

The short mean administration time (6.9 minutes) and low rate of missing values (mean=0.8%) supported the use of the Transdiagnostic Decision Tool in busy clinical settings. In the total sample, all Krippendorff's alpha values exceeded the recommended reliability level of 0.667 [133], demonstrating acceptable inter-rater reliability. However, the Krippendorff's alpha values of item 3 ("Somatic comorbidity") and item 4 ("Psychosocial dysfunctioning") fell short of the recommended reliability level in the SSD subsample. Analyses of the qualitative feedback regarding item 3 suggested that the lower Krippendorff's alpha might be due to the differential classification of medically unexplained physical symptoms across items. In other words, in some instances, clinicians may have classified medically unexplained physical symptoms under item 3 ("Somatic comorbidity") instead of under items concerning the primary diagnosis, such as item 1 ("Severity"). The provided qualitative feedback provided no explanation for the lower Krippendorff's alpha of item 4. Future studies should evaluate whether further specification and clarification of scoring instructions of items 3 and 4 could improve the inter-rater reliability of these items in patients with SSD. The Transdiagnostic Decision Tool demonstrated excellent validity, both in the total sample and within each diagnostic group. Specifically, the total Transdiagnostic Decision Tool score demonstrated meaningful patterns of correlations with total HRQoL and wellbeing scores, supporting convergent validity. In addition, the AUC in the total criterion validity sample was 0.81, and a cut-off value of 3 or greater on the Transdiagnostic Decision Tool was found the optimal cut-off value both in the total sample and within each diagnostic group, indicating that the optimal cut-off value is uniform across these diagnostic groups. Hence, the findings of the present study suggest that, while disorder-specific symptoms are the predominant factors defining differential diagnoses, the allocation of patients to highly specialized healthcare may be meaningfully guided by a core set of transdiagnostic patient factors.

This study has a number of strengths, including the large sample size, the population-based design, and the examination of important psychometric properties related to the use of the Transdiagnostic Decision Tool in daily clinical practice. However, several limitations should also be noted. First, in the absence of a reference test for the systematic and standardized early identification of patients with a highly specialized mental healthcare need, the clinical judgement of clinicians was the reference standard for the evaluation of the criterion validity. Although the use of the clinical judgement as the reference standard may have introduced subjective error, effort was made to reduce error by basing the final clinical judgement on dual, independently provided examinations by highly trained clinicians and extensive experience in the treatment of patients with severe and complex mental health problems. Second, since this study presented a first psychometric evaluation of the Transdiagnostic Decision Tool, future studies

are needed to extend these findings. More specifically, future studies are required to evaluate the psychometric properties of the Transdiagnostic Decision Tool in other diagnostic groups, in patients without a clear primary diagnosis, and in other settings such as primary care. In addition, although the validity of the Transdiagnostic Decision Tool approximates the validity of available diagnosis-specific Decision Tools, future studies are needed to determine whether the Transdiagnostic Decision Tool could substitute these available diagnosis-specific Decision Tools for the diagnostic groups personality disorders, eating disorders, unipolar depression and anxiety disorders, OCD and PTSD. Given the time constraints and competing clinical demands of clinicians in daily practice [177], a trade-off should be made between validity (i.e. precision) and feasibility (i.e. ease of use) of application of the Transdiagnostic Decision Tool in all diagnoses groups. Third, in order to enhance the feasibility of the Transdiagnostic Decision Tool, the scoring system of the tool was constructed as a simple, additive, unweighted sum score. Although this enhances the ease of use in daily clinical practice, it potentially masks differences in the relative importance of individual scale items, which may reduce the precision of the measure. Further work is required to establish the effect of the use of a weighted score on the psychometric properties of the Transdiagnostic Decision Tool. Fourth, notwithstanding its favourable validity in this first study, the Transdiagnostic Decision Tool is intended to augment rather than replace the clinical decision-making process in the referral of patients with mental health problems to highly specialized care. The Transdiagnostic Decision Tool has the potential to provide indications of highly specialized care need, which, together with an assessment of the patient's individual circumstances, preferences and level of motivation, could motivate a referral to treatment in a highly specialized mental healthcare setting. Fifth, although aim of the development of the Transdiagnostic Decision Tool was to facilitate the provision of matched care, the benefit of matched care in patients with a highly specialized mental healthcare is has yet to be studied. Use of the Transdiagnostic Decision Tool in daily clinical practice could, however, enhance the assessment of the clinical- and cost-effectiveness of matched care in patients with a highly specialized mental healthcare need. Finally, although the Transdiagnostic Decision Tool was evaluated for its psychometric properties in specialized and highly specialized mental healthcare settings, the Transdiagnostic Decision Tool might also be of value in primary care services. Use of the Transdiagnostic Decision Tool in primary care services may further enhance the early identification and timely referral of patients with mental health problems in need of highly specialized care. Future studies are required to evaluate the benefit of use of the Transdiagnostic Decision Tool in primary care services.

Despite the limitations, the perceived ease of use, favourable psychometric properties and the transdiagnostic applicability indicate that the Transdiagnostic Decision

Tool can be a promising tool for the early identification and adequate management of patients with mental health problems in need of highly specialized care. Its use in daily practice could enhance the systematic and standardized early identification of patients with a highly specialized mental healthcare need, and thereby has the potential to enhance treatment outcomes, reduce recidivism, reduce prolonged quality of life losses and improve the cost-effective use of scarce healthcare resources.

Appendix 6.1 – Primary items of the diagnosis-specific Decision Tools

Table 6.1.1. Primary items of the diagnosis-specific Decision Tools matching the "Severe primary diagnosis" criterium.

Decision Tool	Item
Eating disorders	Various items that give an indication of the severity of the eating disorder, such as: – Extreme low BMI or extreme high BMI – Duration of the disorder ≥ 2 years
Personality disorders	Severity of the personality disorder
Unipolar depression	Does the patient have severe depression?
Anxiety disorders	Does the patient have a severe anxiety disorder, OCD, or PTSD?

Table 6.1.2. Primary items of the diagnosis-specific Decision Tools matching the "Treatment-interfering psychiatric comorbidity"criterium.

Decision Tool	Item
Eating disorders	Presence of two of more axis I or II comorbidities that interfere with the eating disorder or the treatment of the eating disorder
Personality disorders	Axis I and/or II comorbidity
Unipolar depression	Does the patient have other severe psychiatric comorbidity that interferes with the depression treatment?
Anxiety disorders	Does the patient have at least one diagnosed comorbid psychiatric disorder that interferes with the anxiety, OCD, or PTSD treatment?

Table 6.1.3. Primary items of the diagnosis-specific Decision Tools matching the "Treatment-interfering somatic comorbidity" criterium.

Decision Tool	Item
Eating disorders	Presence of somatic comorbidity
Personality disorders	-
Unipolar depression	Does the patient have somatic comorbidity that interferes with the depression treatment?
Anxiety disorders	Poor physical health

Table 6.1.4. Primary items of the diagnosis-specific Decision Tools matching the "Treatment-interfering psychosocial dysfunctioning" criterium.

Decision Tool	Item
Eating disorders	-
Personality disorders	Severe psychosocial dysfunctioning (GAF≤50)
Unipolar depression	Does the patient exhibit severe psychosocial dysfunctioning?
Anxiety disorders	Does the patient exhibit severe psychosocial dysfunctioning? that interferes with the anxiety, OCD, or PTSD treatment?

Table 6.1.5. Primary items of the diagnosis-specific Decision Tools matching the "Severe or longstanding childhood trauma" criterium.

Decision Tool	Item
Eating disorders	-
Personality disorders	Severe chronic childhood trauma
Unipolar depression	Does the patient have a history of prolonged trauma/neglect in childhood?
Anxiety disorders	-

Table 6.1.6. Primary items of the diagnosis-specific Decision Tools matching the "Previous unsuccessful treatment of the current primary diagnosis in specialized care" criterium.

Decision Tool	Item
Eating disorders	Previous unsuccessful treatments in specialized mental healthcare
Personality disorders	Previous unsuccessful treatments in specialized mental healthcare
Unipolar depression	In the index episode, have there been any unsuccessful previous treatments in specialized mental health care and does the patient have a - recurrent (more than 2 episodes in the past 5 years) or - chronic (>2 years) course of depression?
Anxiety isorders	Have there been any unsuccessful treatments in specialized mental healthcare for the primary diagnoses?

Appendix 6.2 – Transdiagnostic Decision Tool

Figure 6.2.1. English translation of the full and final Transdiagnostic Decision Tool.

Transdiagnostic Decision Tool

Name of patient:

Date:

Name of clinician:

		Comment
1. Does the patient have a severe primary diagnosis? Note: • The primary diagnosis is the most resource-intensive diagnosis. • Indication of severe: SCL-90 high or very high in comparison to a normative sample of outpatients; BSI high or very high in comparison to a normative sample of outpatients; DSM-5 severe.	○ yes ○ no	
2. Does the patient have at least one diagnosed comorbid psychiatric disorder that interferes with the treatment of the primary diagnosis? Note: • Also think of personality disorders, development disorders, addiction, and intellectual disabilities. • There is interference if the diagnosed comorbid disorder complicates the clinical presentation of the primary diagnosis, or the treatment of the primary diagnosis.	○ yes ○ no	
3. Does the patient have somatic comorbidity that interferes with the treatment of the primary diagnosis? Note: There is interference if the somatic comorbidity complicates the clinical presentation of the primary diagnosis, or the treatment of the primary diagnosis.	○ yes ○ no	
4. Does the patient exhibit severe psychosocial dysfunctioning that interferes with the treatment of the primary diagnosis? Note: • Indication of severe psychosocial dysfunctioning: GAF≤50 or WHODAS≥130. • There is interference if the psychosocial dysfunctioning complicates the clinical presentation of the primary diagnosis, or the treatment of the primary diagnosis.	○ yes ○ no	
5. Does the patient have a history of prolonged trauma/neglect in childhood?	○ yes ○ no	
6. Have there been any unsuccessful evidence-based treatments in specialized mental healthcare for the primary diagnosis?	○ yes ○ no	

Total amount of positive (=yes) scores ≥ 3?

Yes → Indicated for highly specialized care on the basis of the Decision Tool

No → Not indicated for highly specialized care on the basis of the Decision Tool

 This work was funded by the Innovation Fund Health Insurers (in Dutch "Innovatiefonds Zorgverzekeraars", https://www.innovatiefondszorgverzekeraars.nl). The Transdiagnostic Decision Tool has been released for use and may be copied. It is not allowed to change the text. For questions or for reporting experiences with the decision tool, please email us at info@topggz.nl.

7

Instruments to assess quality of life in people with mental health problems: A systematic review and dimension analysis of generic, domain- and disease-specific instruments

van Krugten FCW
Feskens K
Busschbach JJV
Hakkaart-van Roijen L
Brouwer WBF

Health and Quality of Life Outcomes. 2021:19(249).

Abstract

Objectives

The importance of economic evaluations of mental healthcare interventions is increasingly recognized. Despite the multitude of available quality of life instruments, concerns have been raised regarding the content validity of these instruments, and hence suitability for use in mental health. The aim of this paper, therefore, was to assess the content validity and suitability of existing quality of life instruments for use in economic evaluations in the mental health field.

Methods

In order to identify available quality of life instruments used in people with mental health problems, a systematic review was performed on the Embase, Medline and PsycINFO databases (time period January 2012 to January 2018). Two reviewers independently assessed study eligibility and executed data extraction. The evaluation framework of Connell et al. was used to assess whether the identified quality of life instruments cover the dimensions valued highly by people with mental health problems. Two reviewers independently mapped the content of each identified instrument onto the evaluation framework and indicated the extent to which the instrument covered each of the dimensions of the evaluation framework.

Results

Searches of databases yielded a total of 5,727 references. Following duplicate removal and double-independent screening, 949 studies were included in the qualitative synthesis. A total of 44 unique quality of life instruments were identified, of which 12 were adapted versions of original instruments. The best coverage of the dimensions of the evaluation framework of Connell et al. was by the WHOQOL-100, S-QoL, SQLS, EDQoL, QLI and the IMHQOL, but none fully covered all dimensions of the evaluation framework.

Conclusions

The results of this study highlight the multitude of available quality of life instruments used in people with mental health problems and indicate that none of the available quality of life instruments fully cover the dimensions previously found to be important in people with mental health problems. Future research should explore the possibilities of refining or expanding existing instruments as well as the development and testing of new quality of life instruments to ensure that all relevant quality of life dimensions for people with mental health problems are covered in evaluations.

Introduction

In the context of scarce resources and rising demands for healthcare, the importance of economic evaluations of healthcare interventions to aid decision makers in allocating healthcare resources is increasingly recognized [178, 179]. Such compare the costs and benefits of healthcare interventions, relative to a relevant comparator, in order to assess their value for money. While costs are typically expressed in monetary terms in such evaluations, benefits are usually expressed in quality-adjusted life-years (QALYs). QALYs comprise changes in both length and quality of life, with the latter typically being measured by generic health-related quality of life instruments, which facilitates comparisons across conditions and interventions [180]. Given the importance of quality of life measurement and valuation in economic evaluations, it is vital to ensure that the instruments used are comprehensive and psychometrically sound.

In the mental health field, the need to assess the relative value for money of different interventions, to inform healthcare resource allocation decisions at different levels, has also been recognised. However, in that context, there is an ongoing debate about how and with which instruments the benefits of mental healthcare interventions could be adequately measured and valued [37, 38]. This topic is particularly relevant for mental health interventions, since alleviating symptoms and improving quality of life are common goals of mental health interventions, rather than prolonging length of life. The adequacy of often used generic health-related quality of life instruments, such as the EuroQol five-dimensional (EQ-5D) questionnaire [181] and the 36-item Short-Form Health Survey (SF-36) [182], has been questioned in the context of (parts of) mental healthcare [37, 38]. More specifically, some have suggested that these instruments, in certain situations, lack the sensitivity to sufficiently reflect the impact of mental health problems on quality of life [38]. The EQ-5D, for example, appears to perform well in mild to moderate mental health conditions [183, 184], but showed weak correlations with severe mental health problems such as schizophrenia [37]. Some argue that this may be due to the fact that these commonly used quality of life instruments have been developed top-down by clinicians or other experts and primarily for people with a physical illness, thereby limiting the coverage of dimensions perceived important to the quality of life of people with mental health problems [39]. Hence, the debate in this area relates both to the sensitivity of existing health-related quality of life instruments, but also to the scope of relevant outcomes (i.e. potentially broadening the evaluative space). The latter is analogous to discussions related to outcome measurements in economic evaluations in elderly care [185]. Another explanation could be that generic instruments by definition focus on the most important quality of life dimensions across diseases, and hence may focus less on particular dimensions relevant in specific dis-

eases. This highlights the tension between the use of generic instruments and more domain or disease specific instruments, which is characterized by a trade-off between comparability between diseases and sensitivity within a disease.

In order to adequately measure and value the benefits of mental healthcare interventions, the use of a multidimensional, preference-based instrument that comprehensively captures the benefits of mental healthcare interventions is required. Based on previous work by Connell et al. [40, 41] that identified seven dimensions known to be important to the quality of life of people with mental health problems, the aim of this paper was to assess the content validity of quality of life instruments used in the mental health field. In addition, it was evaluated whether the available instruments are suitable or, on the basis of the content validity, can be made suitable for use in economic evaluations. The results of this study may then enhance the selection of the most suitable instruments in terms of their coverage of dimensions and benefit the development of adequate outcome instruments to measure and value the benefits of mental healthcare interventions

Methods

Data sources and search strategy

In order to identify available quality of life instruments used in people with mental health problems, a systematic literature search was conducted on the Embase, Medline and PsycINFO databases in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [65]. The search was conducted on January 3, 2018 and was restricted to studies published between January 1, 2012, to January 3, 2018. The search strategy combined terms related to quality of life (e.g. 'quality of life', 'quality of life assessment') and terms related to a broad range of clinical and subclinical mental health problems. See Appendix 7.1 for the search strategies. We did not register a protocol for the review.

Eligibility criteria

Studies were selected for inclusion if they met all of the following criteria:

1. The study population consisted of patients 18 years or older with a clinical or subclinical primary mental health problem;
2. Quality of life was an explicit outcome measure;
3. Quality of life was measured as a multidimensional construct through a generic, domain (i.e. mental health), or disease-specific quality of life instrument with established psychometric properties;

4. The study was a randomized controlled trial, case-control study, cross-sectional study, or cohort study;
5. Published in English and full text available.

Exclusion was based on not meeting all eligibility criteria. Hence, studies that did not meet one or more of the above-listed eligibility criteria were excluded from the review. We emphasise that our review was not restricted to preference-based instruments, but also included 'non-preference-based' instruments. For preference-based or preference-accompanied instruments a value set is available of 'utility scores' that reflect the relative importance of or preference for the states described with such instruments. Such 'utility scores' are typically obtained in a representative sample of the general population and, if derived appropriately, enable the generation of health state utility values for the states described with the instrument. Health state utility values are used to calculate QALYs in economic evaluations of (mental) healthcare interventions. The outcomes of such evaluations can be used in funding and allocation decisions in healthcare. The most frequently used preference-based instrument is the EQ-5D [181]. Other well-known preference-based instruments are the SF-36 [182] and the World Health Organization Quality of Life questionnaire [191].

Study selection and data abstraction

Search results were compiled and deduplicated using RefWorks (http://www.refworks.com), a web-based, bibliographic citation manager. Prior to the eligibility assessment of all identified references, two reviewers independently screened a random sample of 166 titles and abstracts, and reached strong agreement (Cohen's κ=0.83). Blinded to journal titles and authors, the two reviewers then independently screened titles and abstracts of all identified references for potential eligibility using a standardized Excel workbook [66]. For all references that were potentially eligible, a full-text version was retrieved and independently assessed by the reviewers. Disagreements were resolved by discussion or through third-party adjudication. Data abstraction was performed in duplicate and independently using a standardized, Excel-based data abstraction form. The following data were extracted from the included studies: 1) general study characteristics (year of publication, continent of study origin); 2) sample size; 3) (sub) clinical diagnosis of study population; 4) quality of life instrument(s) used. Following the data abstraction of included studies, the development papers and original instruments of identified quality of life instruments were retrieved online or requested from the author. A detailed risk of bias assessment of the included studies was not performed as the primary objective of the review was to compile a list of quality of life instruments used in people with mental health problems.

Evaluation of identified instruments

The following aspects of each of the identified instruments were evaluated: 1) type (generic, domain, or disease-/subgroup-specific); 2) number of items; 3) number of dimensions; 4) region of development; 5) availability of preferences weights (yes/no). The availability of preferences weights was evaluated in order to assess the instruments' suitability for use in cost-effectiveness studies. Such preference-based weights may also help to get an idea about the relative importance of (changes in) different domains and levels. Adapted versions of original instruments were analysed separately as their number of items as well as their number and type of dimensions covered could differ from the original instrument.

An evaluation framework of dimensions was established in order to assess whether the identified quality of life instruments cover the dimensions valued highly by people with mental health problems. The evaluation framework was established based on previous work of Connell et al. [40, 41] that identified seven dimensions known to be important elements of the quality of life of people with mental health problems: well-being and ill-being; relationships and belonging; activity; self-perception; autonomy; hope and hopelessness; physical health. The work by Connell et al. [40, 41] was selected as the basis for the evaluation framework given that it specifically aimed to identify the dimensions of quality of life important to people with mental health problems by using a rigorous mixed-methods approach, i.e. combining a systematic review of qualitative research [40] with complementary interviews [41].

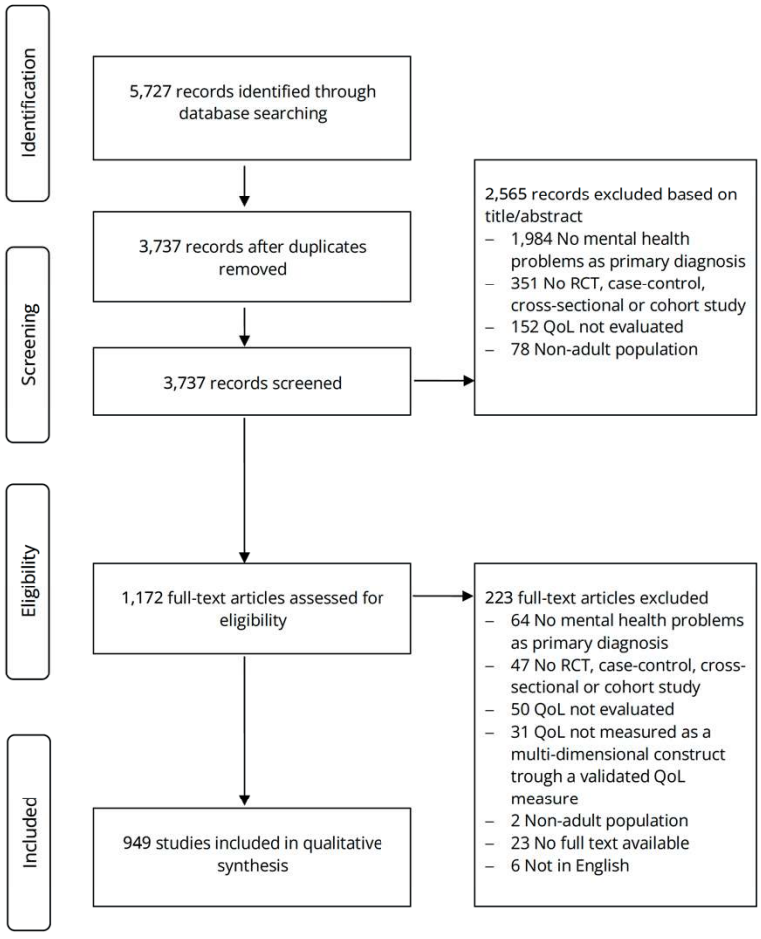
Two reviewers independently mapped the content of each quality of life instrument onto the evaluation framework and indicated the extent (fully, partially, not) to which the instrument covered each of the dimensions of the evaluation framework. A dimension of the evaluation framework was scored as 'fully covered' when the content of the identified quality of life instrument covered more than 75% of the underlying themes of a dimension of the evaluation framework of Connell et al. [40, 41]. Likewise, a dimension was scored as 'partially covered' when the dimensions covered less than 75% of the underlying themes of the dimensions of the evaluation framework. A dimension was scored as 'not covered' when the dimensions covered none of the underlying themes of the dimensions of the evaluation framework. Disagreements were resolved by discussion or through third-party adjudication.

Results

Study selection and study characteristics

The primary search of databases yielded 5,727 references. After duplicate removal and subsequent title and abstract screening, 1,172 papers were obtained for full-text review. Following full-text review, 949 studies met the inclusion criteria and were included in the qualitative synthesis. See Figure 7.1 for the flow chart of the study selection process. The reference list of the 949 included studies is available upon request from the author.

Figure 7.1. Flow chart of study selection process.



RCT = Randomized controlled trial; QoL= Quality of Life.

An overview of the general characteristics of the included studies is shown in Table 7.1. Most of the studies were conducted in Europe (35.6%), followed North America (24.9%), and Asia (23.0%). The most frequently studied diagnosis was schizophrenia spectrum and other psychotic disorders, which was the primary diagnosis of the patient population in 31.3% of the included studies.

Table 7.1. General characteristics of the included studies (N=949).

	N	%
Year of publication		
2012	140	14.8
2013	158	16.6
2014	132	13.9
2015	188	19.8
2016	184	19.4
2017	147	15.5
Study region		
Africa	13	1.4
Asia	218	23.0
Oceania	39	4.1
Europe	338	35.6
Middle east	49	5.2
North America	236	24.9
South America	53	5.6
Covering various regions	3	0.3
Mental health disorder(s) of the study population[a,b]		
Anxiety disorders	41	4.3
Bipolar or related disorders	53	5.6
Depressive disorders	128	13.5
Feeding and eating disorders	42	4.4
Gender dysphoria	2	0.2
Neurodevelopmental disorders	23	2.4
Obsessive-compulsive and related disorders	25	2.6
Personality disorders	4	0.4
Schizophrenia spectrum and other psychotic disorders	297	31.3
Sleep-wake disorders	5	0.5
Somatic symptom and related disorders	7	0.7
Substance-related and addictive disorders	101	10.6
Trauma and stressor-related disorders	55	5.8
Various disorders	166	17.5

[a] Disorders were grouped according to the Diagnostic and Statistical Manual (DSM)-5 [61] categories.
[b] Disorder in this case refers to both clinical and subclinical mental health disorders.

Characteristics of identified instruments

A total of 44 quality of life instruments were identified in the primary search, of which 12 were adapted versions of original instruments. Of all instruments, the World Health Organization's Quality of Life Instrument-Short Version (WHOQOL-BREF) [186] was used most frequently (n=240, 23.9%), followed by the 36-item Short-Form Health Survey (SF-36) [182] (n=181, 18.0%). Of the 44 identified instruments, sixteen were generic instruments, nine were domain-specific (i.e. mental health specific) instruments and nineteen were disease- or subgroup-specific instruments. Generic instruments were the most commonly used (65.0%), followed by domain-specific instruments (20.3%), and disease- and subgroup-specific instruments (14.7%). Of the disease- and subgroup-specific instruments, six were developed for schizophrenia, five for eating disorders, two for veterans, while one was developed for attention-deficit/hyperactivity disorder (ADHD), bipolar depression, Gilles de la Tourette syndrome, forensic inpatients, older people, and patients under neuroleptic treatments. On average, the identified instruments included 35 items (median=23, range=5-143), and covered and average of seven dimensions (median=7, range=2-17). Five instruments allowed for utility score calculations: the Short Form-6 Dimensions (SF-6D) [187], the EuroQol five-dimensional questionnaire (EQ-5D) [181], the Assessment of Quality of life-4 Dimensions (AQoL-4D) [188], the Assessment of Quality of life-8 Dimensions (AQoL-8D) [189], and the 15 Dimensional (15D) [190]. See Table 7.2 for the general characteristics of the identified instruments and Table 7.4 to 7.6 for the complete list of identified quality of life instruments, their frequency of use, and number of items and dimensions.

Table 7.2. General characteristics of the identified instruments.

	All instruments (N=44)	**Generic instruments (N=16)**	**Domain-specific instruments (N=9)**	**Disease- and subgroup-specific instruments (N=19)**
Frequency of use (N, %)	1,004 (100.0)	653 (65.0)	204 (20.3)	147 (14.7)
Number of items				
Mean (SD)	37.2 (35.5)	23.3 (25.8)	66.2 (50.7)	35.1 (26.8)
Median	23	14	78	27
Range	4-143	4-100	15-143	12-131
Number of dimensions				
Mean (SD)	7.6 (3.8)	6.1 (3.0)	10.8 (3.7)	7.4 (3.7)
Median	8	6	9	8
Range	2-17	3-15	8-17	2-15
Number of adapted versions (N, %)	12 (27.3)	9 (56.3)	1 (11.1)	4 (21.1)
Utility score available, Yes (N, %)	5 (11.4)	5 (31.3)	0 (0.0)	0 (0.0)

SD = Standard Deviation.

Instruments' coverage of dimensions of the evaluation framework

The identified instruments differed in the extent to which they covered the dimensions of the evaluation framework (Table 7.3 to 7.6). The "Relationships and belonging" dimension was the most frequently covered (93%), followed by the "Activity" (89%), and "Physical health" (86%) dimensions. The least covered dimensions were the "Self-perception" and "Hope and hopelessness" dimensions, which were included in only 57% and 32% of all instruments, respectively. Compared to the generic instruments and disease- and subgroup-specific instruments, the quality of life instruments specially designed for use in the mental health field covered the "Relationships and belonging", "Activity", and "Autonomy" dimensions most frequently. Of all identified instruments, the World Health Organization's Quality of Life Instrument (WHOQOL-100) [191], the Schizophrenia Quality of Life Questionnaire 41 (S-QoL 41) [192], the Schizophrenia Quality of Life Scale (SQLS) [193], the Eating Disorder Quality of Life (EDQoL) [194], the Quality of Life Index (QLI) [195] and the Internet Mental Health Quality of Life scale (IMHQOL) [196] covered the dimensions of the evaluation framework best. None of the identified instruments fully covered all dimensions of the evaluation framework.

Table 7.3. Frequency (N (%)) with which the identified quality of life instruments (fully or partially) cover the dimensions of the evaluation framework.

Quality of life dimension	All instruments (N=44)	Generic instruments (N=16)	Domain-specific instruments (N=9)	Disease- and subgroup-specific instruments (N=19)
Well-being and ill-being	36 (82)	13 (81)	6 (67)	17 (89)
Relationships and belonging	41 (93)	13 (81)	9 (100)	19 (100)
Activity	39 (89)	13 (81)	9 (100)	17 (89)
Self-perception	25 (57)	8 (50)	5 (56)	12 (63)
Autonomy	28 (64)	7 (44)	9 (100)	12 (63)
Hope and hopelessness	14 (32)	3 (19)	2 (22)	9 (47)
Physical health	38 (86)	16 (100)	9 (100)	13 (68)

Discussion

The aim of this systematic review was to assess the content validity and suitability for use in economic evaluations of quality of life instruments used in people with mental health problems. A total of 44 unique instruments were identified, of which 12 were adapted versions of original instruments. The evaluation framework of Connell et al. [40,41] was used to assess whether the identified quality of life instruments cover the dimensions valued highly by people with mental health problems. The best coverage of the dimensions of the evaluation framework was by the WHOQOL-100, S-QoL 41,

Table 7.4. General characteristics and dimension coverage of the identified generic instruments.

		General characteristics				Dimension coverage[a]						
		N items	*N dimensions*	*Preference based*	*Frequency of use (%)*	*Well-being and ill-being*	*Relationships and belonging*	*Activity*	*Self-perception*	*Autonomy*	*Hope and hopelessness*	*Physical health*
1.	Short-Form Health Survey											
	SF-36 [182]	36	8	N	18.0	X	X	X				X
	SF-12 [197]	12	8	N	10.6	X	X	X				X
	SF-6D[b] [187]	11	6	Y	0.3	X	X	X				X
2.	World Health Organization Quality of Life questionnaire											
	WHOQOL-100 [191]	100	6	N	0.5	X	X	X	X	X	X	X
	WHOQOL-BREF [186]	26	4	N	23.9	X	X	X	X	X	X	X
	EUROHIS-QOL 8-item index [198]	8	4	N	0.4	X	X	X	X	X	X	X
3.	EuroQol five-dimensional questionnaire											
	EQ-5D [181]	5	5	Y	9.0	/		/				X
4.	Assessment of Quality of Life											
	AQoL-4D [188]	12	4	Y	0.4	X	X					X
	AQoL-8D [189]	35	8	Y	0.5	X	X	/	/	/		X
5.	15 Dimensional											
	15D [190]	15	15	Y	0.5	X		/				X
6.	Quality of Life Index											
	QLI [195]	64	4	N	0.5	X	X	X	X	X	X	X
7.	Flanagan's quality of life scale											
	QOLS-15 [199, 200]	15	5	N	0.1		/	/	/	/		X
	QOLS-16 [199, 200]	16	6	N	0.1		/	/	/	/		X
8.	Centers for Disease Control Health-Related Quality of Life Core Module											
	CDC HRQOL-4 [201]	4	3	N	0.1	X		/				X
9.	Personal Wellbeing Index-Adult											
	PWI-A [202]	8	8	N	0.1	/	/			/	/	/
10.	QoL5 [203]	5	3	N	0.1	/	/		/	/		/

Y = Yes; N = No.

[a] X indicates the dimension is fully covered; / indicates the dimension is partially covered.

[b] The SF-6D was the reported instrument in three studies (0.3%); these studies did not report the actual administered instrument (i.e. SF-36 or SF-12).

Table 7.5. General characteristics and dimension coverage of the identified domain-specific instruments.

		General characteristics				Dimension coverage[a]						
		N items	*N dimensions*	*Preference based*	*Frequency of use (%)*	*Well-being and Ill-being*	*Relationships and belonging*	*Activity*	*Self-perception*	*Autonomy*	*Hope and hopelessness*	*Physical health*
1.	Quality of life, Enjoyment, and Satisfaction Questionnaire											
	Q-LES-Q [86]	93	8	N	5.5	X	X	X	X	X		X
	Q-LES-Q-SF [86]	16	8	N	3.9	X	X	X		X		X
2.	Lehmans QoL interview											
	LQLI[b] [204]	143	8	N	3.4		/	X		/		X
3.	Manchester Short Assessment of Quality of Life											
	MANSA [205]	16	8	N	3.3	X	X	X		/		X
4.	Quality of Life Inventory											
	QOLI [206]	17	17	N	1.9		/	X	/	/	/	X
5.	Lancashire Quality of Life Profile											
	LQoLP [207]	105	9	N	1.2	X	X	X	X	X		X
6.	Wisconsin Quality of Life Index for Mental Health											
	W-QLI [208]	113	9	N	0.7	X	X	X	/	/		X
7.	Satisfaction with Life Domains Scale											
	SLDS [209]	15	15	N	0.4		/	/		/		X
8.	Internet Mental Health Quality of Life scale											
	IMHQOL [196]	78	15	N	0.1	X	X	X	X	X	X	X

Y = Yes; N = No.

[a] X indicates the dimension is fully covered; / indicates the dimension is partially covered.

[b] The mapping of the content onto the evaluation framework was based on the description of the items and dimensions in the development paper, since the instrument itself could not be retrieved (online or from the author).

SQLS, EDQoL, QLI and IMHQOL, but none fully covered the dimensions of the evaluation framework. The instruments with the best coverage of the dimensions of the evaluation framework lack a preference-based scoring algorithm, at present. In line with the study of Touré et al. [226], it was found that all identified preference-based instruments, which were all generic, cover the dimension "Physical health", but generally lack coverage of mental health-related (sub)dimensions. Of the five instruments

Table 7.6. General characteristics and dimension coverage of the identified disease- and subgroup-specific instruments.

		General characteristics					Dimension coverage[a]						
		N items	*N dimensions*	*Preference based*	*Frequency of use (%)*	*Disease or subgroup*	*Well-being and ill-being*	*Relationships and belonging*	*Activity*	*Self-perception*	*Autonomy*	*Hope and hopelessness*	*Physical health*
1.	Heinrichs-Carpenter Quality of Life Scale												
	QLS[b] [210]	21	4	N	4.5	SCZ	✓	✓	✓	✓			
2.	Schizophrenia Quality of Life Questionnaire												
	S-QoL 41 [192]	41	8	N	0.7	SCZ	✓	✓	✓	✓	✓	✓	✓
	S-QoL 18 [211]	18	8	N	2.3	SCZ	✓	✓	✓	✓	✓	✓	✓
3.	Schizophrenia Quality of Life Scale												
	SQLS [193]	30	3	N	1.9	SCZ	✓	✓	✓	✓	✓	✓	✓
4.	Eating Disorder Quality of Life												
	EDQoL [194]	25	4	N	0.7	ED	✓	✓	✓	✓	✓	✓	✓
5.	Adult ADHD Quality of Life Scale												
	AAQol [212]	29	4	N	0.7	ADHD	✓	✓	✓	✓	✓		
6.	Veterans rand health survey												
	VR-36 [213]	36	8	N	0.4	VT	✓	✓	✓				✓
	VR-12 [214]	12	8	N	0.2	VT	✓	✓	✓				✓
7.	Quality-of-Life in Schizophrenia												
	QLiS[b] [215]	52	12	N	0.6	SCZ	✓	✓	✓	✓	✓		
8.	Brief version of Quality of Life in Bipolar Disorder												
	Bref QoL.BD [216]	12	12	N	0.5	BP	✓	✓	✓		✓		✓
9.	Seville Quality of Life Questionnaire												
	CSCV [217]	59	12	N	0.5	SCZ	✓	✓	✓	✓	✓	✓	
10.	Forensic inpatient Quality of Life questionnaire												
	FQL [218]	131	15	N	0.4	FI		✓	✓	✓	✓	✓	✓
11.	Health-Related Quality of Life in Eating Disorders												
	HeRQoLED [219]	50	8	N	0.2	ED	✓	✓	✓	✓			✓
	HeRQoLED-s [220]	20	2	N	0.2	ED	✓	✓	✓	✓			

Table 7.6. General characteristics and dimension coverage of the identified disease- and subgroup-specific instruments. (continued)

		General characteristics					Dimension coverage[a]							
		N items	*N dimensions*	*Preference based*	*Frequency of use (%)*	*Disease or subgroup*	*Well-being and ill-being*	*Relationships and belonging*	*Activity*	*Self-perception*	*Autonomy*	*Hope and hopelessness*	*Physical health*
12.	Subjective Well-Being Under Neuroleptic Treatment Scale short form												
	SWN-20 [221]	20	5	N	0.4	PUNT	X	X		/	/	/	
13.	World Health Organization Quality of Life Questionnaire-Older Adults Module												
	WHOQOL-OLD [222]	24	6	N	0.2	Elderly		X	X		/	X	X
14.	Eating Disorders Quality of Life Survey												
	EDQLS [223]	40	12	N	0.1	ED	X	X	X	X		X	X
15.	Gilles de la Tourette syndrome-Quality of life scale												
	GTS-QoL [224]	27	4	N	0.1	GTS	X	/	/	/	/		X
16.	Quality of Life Eating Disorders												
	QOL ED[b] [225]	20	5	N	0.1	ED	/	/	/				X

ADHD = Attention-Deficit/Hyperactivity Disorder; BP = Bipolar Depression; ED = Eating Disorders; FI = Forensic Inpatients; GTS = Gilles de la Tourette; N = No; PUNT = Patients Under Neuroleptic Treatment; SCZ = Schizophrenia; Y = Yes; VT = Veterans.

[a] X indicates the dimension is fully covered; / indicates the dimension is partially covered.

[b] The mapping of the content onto the evaluation framework was based on the description of the items and dimensions in the development paper, since the instrument itself could not be retrieved (online or from the author).

that were found to have a preference-based scoring algorithm, the AQoL-8D had the most overlap with the framework of Connell et al. [40, 41].

The results of this study highlight the multitude of available quality of life instruments and support previous research questioning the ability of commonly used instruments to adequately measure and value the benefits of mental healthcare interventions [39]. The findings of this review suggest that this inability might be related to the content validity of the available quality of life instruments, since none of the identified preference-based instruments was found to fully cover the dimensions known valued highly by people with mental health problems. Noteworthy was the lack of coverage of the "Hope and hopelessness" and "Self-perception" dimensions, which were covered in

only 14% and 34% of the identified instruments, respectively. Note that the low coverage of the "Hope and Hopelessness" dimension may be explained by the fact that this dimension may be, to a certain degree, transversal to depression and distress, which were underlying themes of the "Well-being" dimension. It is important to recognize differences in the coverage of dimensions in selecting the quality of life instruments of choice for evaluating the effectiveness of interventions, as they implicitly define the maximand of interventions. Another noteworthy finding was that the majority of identified instruments are non-preference-based and are, therefore, not directly useful for inclusion in cost-utility studies. In order to make available instruments suitable for use in cost-utility studies, health state utility values should be generated by use of utility-elicitation procedures or, as a second-best option, predicted by statistical association [227]. However, given that none of the identified instruments fully cover the dimensions valued highly by people with mental health problems, it seems advisable to first refine existing instruments or develop new quality of life instruments that cover all of the relevant dimensions. In the refinement or development of such instruments, next to their content validity, other elements of validity and reliability require much attention. Even more so, as, particularly in the mental health field, self-completion instruments may be less reliable in certain disease areas and may be prone to bias due to effects of social desirability and stigma. In addition, in order to sufficiently reflect the impact of mental health problems on quality of life, but simultaneously prevent a loss of comparability of utility values across mental health diagnoses, such new instruments should preferably be domain-specific (i.e. mental health) in nature. It needs noting that such a strategy does raise numerous questions about the desired scope of such instruments and the subsequent comparability of outcomes across sectors. In other words, optimization per domain may compromise the optimization over domains. These issues are beyond the scope of the current review but require attention in future research.

This systematic review is strengthened by its use of a comprehensive search strategy, the bias protection measures taken (e.g. independent and duplicate screening and reviewing of identified studies), the executed dimension analysis of identified instruments based on a scientifically founded evaluation framework, and the inclusion of studies focusing on populations with clinical and subclinical primary mental health problems. However, despite the strengths of this review, some limitations should be noted. First, the review was restricted to peer-reviewed studies published in the Embase, Medline and PsycINFO databases. Expanding the search strategy by, for instance, including grey literature, using snowballing or including other databases such as the Cochrane Central Register of Controlled Trials, might have produced (even) more results. Hence, some relevant studies may have been missed in the current review. Second, most of the included studies were conducted in Europe, North America, and Asia. Future

research could explore the reasons for the relatively low frequency of use of quality of life instruments in mental health research in other continents. Third, given our focus on published studies up to 2018, we may have missed recent developments in the field of quality of life assessment. One important quality of life instrument, specifically designed for use in the mental health field, that has become available since the completion of our review is the Recovering Quality of Life (ReQoL) measure [228]. The ReQoL measure is a preference-based [229] patient reported outcome measure that was explicitly designed to cover all seven dimensions of the evaluation framework used in the current study. The development of the ReQoL measure highlights the need and search for outcome instruments that adequately measure and value the benefits of mental healthcare interventions. Further work is required to assess how the ReQoL performs in various contexts, especially in contexts in which existing quality of life measures lack the sensitivity to sufficiently reflect the impact of mental health problems on quality of life, and in relation to other outcome measures identified in this study. Fourth, given the focus on the identification of quality of life instruments used in people with mental health problems, we might have missed relatively new instruments that were available but not used in studies published in the reference period of our search. The Clinical Outcomes in Routine Evaluation (CORE)-6D [230] is an example of such a measure. Given the rapid developments in this field, it is advisable that studies like the present one are repeated in the future. Fifth, since the aim of the review was to identify available quality of life instruments used in people with mental health problems and assess whether these instruments cover the dimensions found to be important in people with mental health problems (content validity), the analysis does not take anything regarding the other psychometric performance of the identified instruments into account. Inclusion of quality of life instruments in studies on the (cost-)effectiveness of mental health interventions should be based on and motivated by evidence on all psychometric properties of the instruments, as for example assessed by the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) [231]. Hence, even if instruments cover most of the dimensions of the evaluation framework, it does not imply that these instruments are recommended over others, nor does it imply that these instruments are the best available for use in people with mental health problems. In addition, failure to meet the criteria of the evaluation framework is not a disqualification of the instrument as such, but it raises questions about the suitability of the instruments when used in the context of mental health. The findings of this study could, however, enhance the selection of the most suitable instruments in terms of their coverage of dimensions and practical characteristics such as number of items and the availability of preference-based utility values. Sixth, the study population of one of the studies underlying the evaluation framework [41] only included mental healthcare service users, not a wider population

of people with mental health problems. This may have influenced the dimensions of quality of life in the framework. However, in the absence of studies examining the important quality of life dimensions in a broader, mixed population with people with mental health problems, the study carried out by Connell et al. [41] was considered the best available to base the framework on. Seventh, the adoption of the framework by Connell et al. [40, 41] implicitly implies that life domains considered important by the relevant population should determine the evaluative scope of an economic evaluation. This matter can be debated and relates to normative questions of what should be maximized (health or more general well-being), whether outcome measures should be generic or may be domain-specific, and the appropriate source for domains and their relative valuations. These are crucial questions that fall outside the scope of the current study. Eighth, the mapping of the dimensions of the identified instruments onto the evaluation framework was inherently subjective. In order to minimise the subjective nature of the mapping procedure, the dimensions of each identified instrument were assessed and mapped onto the evaluation framework by two reviewers in a structured, independent manner using standardized criteria.

The results of this study highlight the multitude of available quality of life instruments and lack of consensus regarding the choice of instruments used in people with mental health problems. In addition, the results could enhance the selection of the most suitable instruments in terms of their coverage of dimensions and practical characteristics. At the same time, the increasing importance of quality of life measurement in clinical and research settings emphasizes the need for more methodological studies on quality of life measurement in the mental health field. More specifically, future research could evaluate and compare the psychometric properties of promising instruments, and obtain preference-based utility values for these instruments to make them suitable for use in cost-effectiveness studies. In addition, since the results of this study suggest that none of the identified instruments cover all the dimensions found to be important in people with mental health problems, future research could explore the possibilities of refining existing instruments or the development of a new quality of life instrument that covers all of the dimensions valued highly by people with mental health problems. Future research on these issues remains crucial to capture the benefits of interventions targeted at people with mental health problems and facilitate the comparison of the clinical and cost-effectiveness of mental healthcare interventions, which in turn could improve the allocation of scarce resources in the mental health field.

Appendix 7.1 – Search strategies

Table 7.1.1. Search strategy Embase.

Search
('quality of life'/mj/exp OR 'quality of life assessment'/mj/exp OR ((quality NEAR/3 life) OR qol):ti) AND ('mental health'/mj/exp OR 'psychiatry'/mj/exp OR 'Diagnostic and Statistical Manual of Mental Disorders'/mj/exp OR 'mental disease'/exp/mj OR 'sexual deviation'/mj OR (((mental* OR psychic*) NEAR/3 (health* OR disorder* OR ill*)) OR (mood NEAR/3 disorder*) OR (obsessi* NEAR/3 compuls*) OR dsm OR psychiatr* OR schizo* OR bipolar OR unipolar OR ((posttrauma* OR trauma*) NEAR/3 stress) OR ptsd OR ((dissociati* OR somatoform* OR eating OR personalit* OR psychosexual* OR behav* OR psychosocial*) NEAR/3 disorder*) OR (multiple NEAR/3 personalit*) OR anorex* OR bulemi* OR addict* OR ((drug* OR alcohol* OR substance*) NEAR/3 (depend* OR misuse OR abuse)) OR borderlin* OR depressi* OR anxi* OR panic OR (Diagnos* NEAR/3 Statistic* NEAR/3 Manual*) OR 'sexual* deviat*' OR paraphil* OR psychosis OR psychoses OR psychotic OR phobia*):ti) NOT ([Conference Abstract]/lim OR [Letter]/lim OR [Note]/lim OR [Editorial]/lim) AND [english]/lim NOT (cancer* OR malign* OR neoplas* OR hiv OR aids OR diabet* OR dement* OR cardiac* OR myocard* OR surg* OR postsurg* OR postoperat* OR stroke* OR cva):ti NOT ((j uvenile/exp OR (child* OR infan* OR adolescen*):ab,ti) NOT (adult/exp OR (adult* OR elderl*):ab,ti))

Table 7.1.2. Search strategy Medline (Ovid).

Search
(* quality of life/ OR ((quality ADJ3 life) OR qol).ti.) AND (* exp mental health/ OR * exp psychiatry/ OR * Diagnostic and Statistical Manual of Mental Disorders/ OR * exp Mental Disorders/ OR (((mental* OR psychic*) ADJ3 (health* OR disorder* OR ill*)) OR (mood ADJ3 disorder*) OR (obsessi* ADJ3 compuls*) OR dsm OR psychiatr* OR schizo* OR bipolar OR unipolar OR ((posttrauma* OR trauma*) ADJ3 stress) OR ptsd OR ((dissociati* OR somatoform* OR eating OR personalit* OR psychosexual* OR behav* OR psychosocial*) ADJ3 disorder*) OR (multiple ADJ3 personalit*) OR anorex* OR bulemi* OR addict* OR ((drug* OR alcohol* OR substance*) ADJ3 (depend* OR misuse OR abuse)) OR borderlin* OR depressi* OR anxi* OR panic OR (Diagnos* ADJ3 Statistic* ADJ3 Manual*) OR sexual* deviat* OR paraphil* OR psychosis OR psychoses OR psychotic OR phobia*).ti.) NOT (letter OR news OR comment OR editorial OR congresses OR abstracts).pt. AND english.la. NOT (cancer* OR malign* OR neoplas* OR hiv OR aids OR diabet* OR dement* OR cardiac* OR myocard* OR surg* OR postsurg* OR postoperat* OR stroke* OR cva).ti. NOT ((exp child/ OR exp infant/ OR adolescent/ OR (child* OR infan* OR adolescen*).ab,ti.) NOT (exp adult/ OR (adult* OR elderl*).ab,ti.))

Table 7.1.3. Search strategy PsycINFO (Ovid).

Search
(* "quality of life"/ OR ((quality ADJ3 life) OR qol).ti.) AND (* exp mental health/ OR * exp psychiatry/ OR * exp Mental Disorders/ OR (((mental* OR psychic*) ADJ3 (health* OR disorder* OR ill*)) OR (mood ADJ3 disorder*) OR (obsessi* ADJ3 compuls*) OR dsm OR psychiatr* OR schizo* OR bipolar OR unipolar OR ((posttrauma* OR trauma*) ADJ3 stress) OR ptsd OR ((dissociati* OR somatoform* OR eating OR personalit* OR psychosexual* OR behav* OR psychosocial*) ADJ3 disorder*) OR (multiple ADJ3 personalit*) OR anorex* OR bulemi* OR addict* OR ((drug* OR alcohol* OR substance*) ADJ3 (depend* OR misuse OR abuse)) OR borderlin* OR depressi* OR anxi* OR panic OR (Diagnos* ADJ3 Statistic* ADJ3 Manual*) OR sexual* deviat* OR paraphil* OR psychosis OR psychoses OR psychotic OR phobia*).ti.) NOT (letter OR news OR comment OR editorial OR congresses OR abstracts).pt. AND english.la. NOT (cancer* OR malign* OR neoplas* OR hiv OR aids OR diabet* OR dement* OR cardiac* OR myocard* OR surg* OR postsurg* OR postoperat* OR stroke* OR cva).ti. NOT ((100.ag. OR 200.ag. OR (child* OR infan* OR adolescen*).ab,ti.) NOT (300.ag. OR (adult* OR elderl*).ab,ti.))

8

The Mental Health Quality of Life questionnaire (MHQoL): Development and first psychometric evaluation of a new measure to assess quality of life in people with mental health problems

van Krugten FCW
Busschbach JJV
Versteegh MM
Hakkaart-van Roijen L
Brouwer WBF

Quality of Life Research. 2021:1-11.

Abstract

Objectives

The aim of this study was to develop and psychometrically evaluate a new quality of life measure for use in people with mental health problems, the Mental Health Quality of Life questionnaire (MHQoL).

Methods

The MHQoL dimensions were based on prior research by Connell et al., highlighting the seven most important quality of life dimensions in the context of mental health. Items were generated following a systematic review we performed and through inviting expert opinion. A focus group and an online qualitative study (N=120) were carried out to assess the face and content validity of the MHQoL. The MHQoL was further tested for its internal consistency, convergent validity, known-group validity and test-retest reliability among mental healthcare service users (N=479) and members of the general population (N=110).

Results

The MHQoL consists of a descriptive system (MHQoL-7D), including seven items covering seven dimensions (self-image, independence, mood, relationships, daily activities, physical health, future) and a visual analogue scale of general psychological well-being (MHQoL-VAS). Internal consistency was high (Cronbach's alpha=0.85) and correlations between MHQoL-7D scores and related measures (EQ-5D-5L, MANSA, ICECAP-A, and BSI) supported convergent validity. The intraclass correlation coefficient of the MHQoL-7D sum score for test-retest reliability was 0.85. Known-group validity was supported by the ability to detect significant differences in MHQoL-7D levels between service users and the general population, and between groups with different levels of psychological distress.

Conclusions

The MHQoL demonstrated favourable psychometric properties and showed promise as a simple and effective measure to assess quality of life in people with mental health problems.

Introduction

The concept of quality of life is widely and increasingly used as an important outcome measure in the evaluation of healthcare interventions [232]. Also in the mental health field, it is recognized that while symptom reduction is a desirable treatment outcome, it is also important to assess how recovery translates to the daily life of an individual and their quality of life [233]. Although a consensual definition is lacking, there is general agreement that quality of life is a subjective and multidimensional construct that captures an individual's life satisfaction and overall well-being [234]. In order to accommodate the growing interest in measuring and monitoring the impact of mental health(care) on peoples' lives, mental healthcare providers in, for example, the Netherlands and the United Kingdom, increasingly include quality of life measures in their routine outcome measurement alongside more clinically oriented measures [235, 236].

Despite the growing interest in assessing quality of life in mental healthcare, it has been questioned whether frequently used quality of life measures, such as the EuroQol five-dimensional (EQ-5D) questionnaire [181] and the 36-item Short-Form Health Survey (SF-36) [182], adequately capture and value the benefits of mental healthcare interventions. Previous studies have indicated that frequently used quality of life measures are, in certain situations, not sufficiently sensitive to the effects of mental health problems on quality of life [38, 39, 183]. It has been argued that this may be due to the large focus on physical health of these commonly used quality of life measures, which limits the coverage of the dimensions of quality of life valued highly by people with mental health problems [39].

A recent systematic review [237; Chapter 7] indicated that the inability of available quality of life measures to adequately capture and value the benefits of mental healthcare interventions might be related to the content validity of these measures. More specifically, it was found that none of the generic (e.g. SF-36 [182]), domain-specific (e.g. Manchester Short Assessment of Quality of Life [205]) or disease-specific (e.g. Schizophrenia Quality of Life Scale [193]) quality of life measures used in people with mental health problems fully cover the dimensions found to be important to the quality of life of people with mental health problems [40, 41]. Those findings underline the need for a measure that covers the dimensions considered to be important by people with mental health problems, providing both a descriptive profile and an overall index.

The present paper reports on the development and psychometric evaluation of the Mental Health Quality of Life questionnaire (MHQoL), designed to comprehensively

provide information about the quality of life dimensions known to be relevant across and valued highly by people with mental health problems. The conceptual framework was established based on previous work carried out by Connell et al. [40, 41]. This work aimed to identify the dimensions of quality of life important to people with mental health problems and has been shown to be an attractive theoretical foundation for the development of quality of life measures for use in the mental health field. Indeed, in the same period in which the MHQoL was developed, Keetharuth et al. developed the Recovering Quality of Life (ReQoL) measures [228], which were also based on this framework. In the discussion section of this chapter, we will reflect on the differences between the MHQoL and the ReQoL measures.

Methods

The study consisted of two major phases: (1) development and (2) psychometric evaluation of the Mental Health Quality of Life questionnaire (MHQoL). The study was reviewed and approved by the Medical Ethical Committee of the Erasmus University Medical Centre Rotterdam, The Netherlands (MEC-2018-142) and digital informed consent was obtained from all participants in the study.

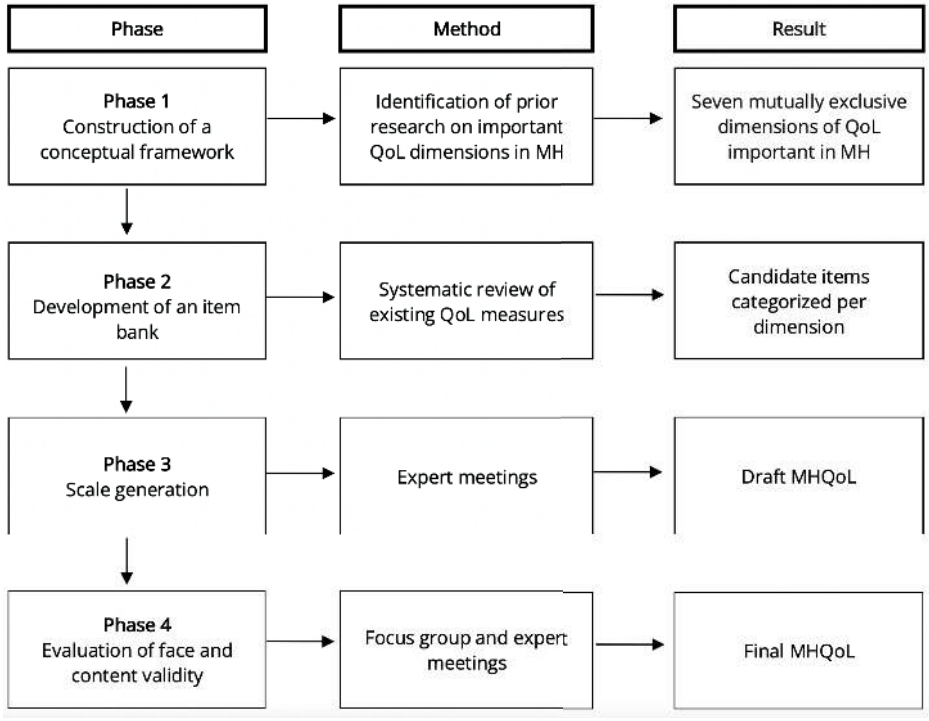
MHQoL development

The first phase of the study, in which the MHQoL was developed, consisted of four stages: (1) construction of a conceptual framework to guide measurement development; (2) development of an item bank to guide item generation; (3) scale generation; and (4) evaluation of face and content validity. See Figure 8.1 for a visual representation and summary of these stages. The development process was led by a group of researchers (n=6) with relevant expertise in the field of scale development, mental healthcare, or in both.

As a first stage in the development process, a conceptual framework was constructed to serve as a theoretical basis for the resultant measure. The conceptual framework was established based on previous work carried out by Connell et al. [40, 41], highlighting seven dimensions of quality of life most important to people with mental health problems (well-being and ill-being; physical health; autonomy; self-perception; relationships and belonging; activity; hope and hopelessness). The work by Connell et al. [40, 41] was selected as the basis for the conceptual framework, given that it specifically aimed to identify the dimensions of quality of life important to people with mental health problems by using a rigorous mixed-methods approach combining a systematic review of qualitative research [40] with complementary interviews [41]. A

visual representation of the dimensions of the conceptual framework can be found in the work by Keetharuth et al. [238]. In the second stage of the development process, a bank of candidate items was developed to inform the generation of MHQoL items. The item bank was developed on the basis of a recent systematic review we performed that aimed to identify existing quality of life measures used in people with mental health problems [237; Chapter 7]. Through examination of the content of the identified measures (n=44), a total of 272 candidate items were extracted and categorized per dimension of the conceptual framework. In three expert meetings, the item bank was reduced by only retaining the items that best covered the underlying themes of the dimensions of the evaluation framework (see Connell et al. [40] for the underlying themes of the dimensions).

Figure 8.1. Development stages of the MHQoL.



QoL = Quality of Life; MH = Mental Health; MHQoL = Mental Health Quality of Life questionnaire.

Informed by the reduced bank of candidate items, preliminary scale items were generated for each of the seven dimensions of the conceptual framework in the third stage of the development process. Main requirements in the generation of items were that the resultant measure should be transdiagnostic in nature and short and easy to complete by the respondent. These principles led to the operationalization of the seven dimen-

sions into seven items (one item per dimension), each with four response options[1]. In line with measures like the EQ-5D and in order to avoid subjective weighting of health states experienced over longer periods of time, the recall period was set to "today". In twelve expert meetings, the generated items were extensively discussed to ensure that all items sufficiently reflected the intended meaning of each of the dimensions. As a result of the discussions, some changes were made to the wording and labels of the items, resulting in the first draft version of the MHQoL.

In the fourth and final stage of the development process, the face and content validity of the draft version of the MHQoL were evaluated in two steps. The first step consisted of a focus group in which six mental healthcare service users were asked to complete the MHQoL, followed by a de-briefing exercise in which they examined the meaning of the individual items, the extent to which the items seem to cover the things that matter in their lives, and the adequacy of the response options. Based on this focus group, minor changes were made to the wording and sequence of the items. In the second stage, a web-based survey was carried out among 120 adult (18 years and older) mental healthcare service users. Participants were randomly drawn from an online panel through the market research company Dynata. Inclusion criteria were: aged 18 years or older and visited any health professional (e.g. psychiatrist, psychologist, general practitioner, social worker) for mental health problems in the past twelve months. Participants were asked to fill out the MHQoL, indicate whether the items cover the things that matter in their lives, and comment on the clarity of the individual items and the measure as a whole. Analysis of the provided comments confirmed the completeness and clarity of the MHQoL; no changes to the wording and sequence of items were deemed necessary.

The Mental Health Quality of Life questionnaire (MHQoL)

The development process resulted in the Mental Health Quality of Life questionnaire (MHQoL). The MHQoL is a standardized, self-administered measure of quality of life that has been developed for use in people with subclinical and clinical mental health problems and across all types of mental health services. The MHQoL consists of two parts: a descriptive system, the MHQoL-7D and a visual analogue scale, the MHQoL-VAS. The MHQoL-7D comprises seven questions, covering seven dimensions (self-image, independence, mood, relationships, daily activities, physical health, future), each with four response levels (e.g. ranging from very satisfied (score=3) to very dis-

1 The item labels of the MHQoL correspond as follows to the labels of the dimensions of the conceptual framework (item label MHQoL = dimension label): self-image = self-perception; independence = autonomy; mood = well-being and ill-being; relationships = relationships and belonging; daily activities = activity; physical health = physical health; future = hope and hopelessness.

satisfied (score=0)). The MHQoL-7D sum score can vary from 0 to 21, with higher scores indicating better quality of life. The MHQoL-VAS records the self-esteemed general psychological well-being of the respondent on a horizontal scale ranging from zero ("worst imaginable psychological well-being") to ten ("best imaginable psychological well-being"). The MHQoL was developed in Dutch. The English translation of the MHQoL is included in Appendix 8.1.

Evaluation of psychometric properties

Study design and population

In order to evaluate the psychometric properties of the MHQoL, a web-based study was carried out. The study population consisted of 479 adult (18 years and older) mental healthcare service users and 110 adult members of the general population. During September 2018, participants were drawn from a consumer panel through the market research company Dynata. The subsample of mental healthcare service users (aged 18 years or older) was selected from the larger panel based on the fact that respondents themselves indicated that they visited any health professional (e.g. psychiatrist, psychologist, general practitioner, social worker) for mental health problems in the past twelve months. The general population subsample was selected to represent the Dutch population in 2018 in terms of the distribution of age and sex as recorded by Statistics Netherlands (Centraal Bureau voor de Statistiek). Participants received a financial incentive of €1.50 for their participation in the study.

Measures

In addition to the MHQoL, participants completed the self-report measures listed below.

- The *five-level EuroQol five-dimensional questionnaire (EQ-5D-5L)* [157] is a five-item generic, preference-based self-report measure to describe and value health-related quality of life (HRQoL). The EQ-5D-5L covers five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and a visual analogue scale (EQ-VAS) for overall health. Each dimension is divided into five response options describing the state per dimension (no problems, some problems, moderate problems, severe problems, and extreme problems/unable to). An index summary score can be generated by applying societal preference weights to the health state classification (scoring on the five dimensions) as completed by the respondent. Based on the Dutch tariff, total scores can range from -0.446 to 1 [158], with higher values indicating better HrQoL as perceived by the general population. The EQ-VAS is a vertical scale ranging from zero ("worst imaginable health state") to 100 ("best imaginable health state") on which the respondents are asked to rate their overall health.
- The *Manchester Short Assessment of quality of life* (MANSA) [205] is a sixteen-item self-report measure to assess quality of life in people with mental health problems.

The MANSA is a shortened version of the Lancashire Quality of Life Profile (LQLP) [207] and consists of four dichotomous (yes/no) items covering objective quality of life aspects and twelve items assessing the satisfaction with life as a whole, job, financial situation, friendships, leisure activities, accommodation, personal safety, people that the person lives with, family and health. Each of the twelve satisfaction items is rated on a seven-point scale ranging from one ("couldn't be worse") to seven ("couldn't be better"). Summary scores can range from 12 to 84, with higher scores indicating better quality of life.

- The *ICEpop CAPability measure for Adults* (ICECAP-A) [162] is a five-item generic, preference-based self-report measure of capability well-being for use in the adult population. The items cover five dimensions (stability, attachment, autonomy, achievement, and enjoyment), and each item has four response levels (e.g. none, a little, a lot and all). Index summary scores can range from 0 (representing the absence of capability) to 1 (representing full capability) [176].
- The *Brief Symptom Inventory* (BSI) [239] is a 53-item self-report measure of psychopathology. The BSI is a shortened version of the Symptom Checklist-90 (SCL-90) [240] and covers nine dimensions (somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism). Each item is rated on a five-point scale ranging from zero ("not at all") to four ("extremely"). The summary scale index of the BSI, the "Global Severity Index" (GSI), can range from 0 to 212, with higher scores indicating greater psychological distress.

Procedures

After providing digital informed consent, participants were asked to complete a web-based survey containing the MHQoL and questions about their socio-demographics (gender, date of birth, level of education, employment/activity) and mental health status (mental health problem, severity of mental health problem, duration of mental health problem). In addition, participants completed the EQ-5D-5L, MANSA, ICECAP-A, and BSI in order to evaluate convergent validity. After one week, the MHQoL was readministered to a randomly selected subset of 33% of participants reporting no change in their mental health-related quality of life status after one week to assess test-retest reliability.

Statistical analysis

Data on the demographic and clinical characteristics of the study sample were analysed using descriptive statistics. Internal consistency was assessed by item-total correlations and Cronbach's alpha coefficient in the total sample and subsample of mental healthcare service users. Cronbach's alpha values of 0.70-0.79 were considered acceptable, 0.80-0.89 good, and ≥0.90 excellent [241]. Test-retest reliability was assessed by intraclass

correlation coefficient (ICC) using the two-way mixed effects, absolute agreement, single measurement model. Intraclass correlation coefficients of <0.49, 0.5-0.74, 0.75-0.89, >0.90 were considered poor, moderate, good, and excellent, respectively [242]. In order to assess convergent validity, Spearman's rank correlations were calculated between total MHQoL-7D scores and EQ-5D-5L, EQ-5D-5L index, EQ-VAS, MANSA, ICECAP-A, ICECAP-A index and BSI scores. Spearman's rank correlations of 0.10-0.29 were considered weak, 0.30-0.49 moderate, and ≥0.50 strong [135]. MHQoL-7D scores were expected to have a strong positive correlation with quality of life (EQ-5D, MANSA) and well-being (ICECAP-A) scores. Since quality of was demonstrated to be sensitive to variations in psychopathology (e.g. [243, 244]), the MHQoL was hypothesized to have a moderate negative correlation with the BSI. Within the subsample of mental healthcare service users, known group validity was assessed by evaluating the ability of the MHQoL-7D to detect significant group differences between participants by clinical status (a clinical BSI score of ≥0.67 vs. a non-clinical BSI score of <0.67 [245]) and self-reported severity of mental health problems (severe vs. mild/moderate). The four original severity categories of mental health problems (mild, moderate, severe, very severe) were collapsed into two categories of (mild/moderate and severe). In addition, known-group validity was assessed in the total sample by testing whether the MHQoL-7D was able to discriminate between mental healthcare users and members from the general population. Group differences were examined using the Mann-Whitney U test. Mean MHQoL-7D group scores were expected to be significantly higher (i.e. better) in the group with non-clinical psychopathology, the group with mild/moderate mental health problems, and in the group of members from the general population. All analyses were carried out using the Statistical Package for the Social Sciences (SPSS) version 24.0 (SPSS Inc., IBM Corporation, Armonk, New York, USA). Significance levels were set at P<0.05 (two-tailed).

Results

Participant's characteristics

Demographic and clinical characteristics of the study sample are presented in Table 8.1. The study sample comprised 479 mental healthcare service users and 110 members of the general population. The mean age of the total sample was 46.5 years (SD=15.8), 341 (57.9%) were female, and most of the participants attained middle education (45.8%). In the subsample of mental healthcare service users, the most commonly reported mental health problems were depression (64.5%), dysthymia (41.8%), and anxiety disorder (42.0%). In the subsample of mental healthcare service users, the mental health problems were, as classified according to the own perception of participants, in most cases of moderate severity (48.4%). The mean total MHQoL-7D

and MHQoL-VAS scores were lower in the subsample of mental healthcare service users (11.5 (4.0) and 5.7 (2.0), respectively) than in the subsample of members of the general population (15.5 (2.9) and 7.5 (1.5), respectively).

Table 8.1. Demographic and clinical characteristics of study sample.

	Total sample	**Mental healthcare service users[a]**	**Members of the general population[a]**
N	589	479	110
Age, years			
Mean (SD)	46.5 (15.8)	46.0 (15.7)	48.6 (16.1)
Range	18.1 - 85.8	18.1 - 85.8	18.9 - 80.7
Sex (N, %)			
Male	245 (41.6)	189 (39.5)	56 (50.9)
Female	341 (57.9)	287 (59.9)	54 (49.1)
Transgender	3 (0.5)	3 (0.6)	0 (0.0)
Education (N, %)[b]			
Lower education	135 (22.9)	113 (23.6)	22 (20.0)
Middle education	270 (45.8)	215 (44.9)	55 (50.0)
Higher education	184 (31.2)	151 (31.5)	33 (30.0)
Visited health professional for mental health problems in past 12 months (N, %)			
Yes	499 (84.7)	479 (100)	20 (18.2)
No	90 (15.3)	0 (0)	90 (81.8)
Type of health professional visited for mental health problems in past 12 months (N, %)[c]			
General practitioner	326 (55.3)	315 (65.8)	11 (10.0)
General nurse Practitioner mental healthcare	136 (23.1)	132 (27.6)	4 (3.6)
Social worker	61 (10.4)	59 (12.3)	2 (1.8)
Occupational physician	43 (7.3)	40 (8.4)	3 (2.7)
Psychotherapist	59 (10.0)	58 (12.1)	1 (0.9)
Psychologist	213 (36.2)	208 (43.4)	5 (4.5)
Psychiatrist	137 (23.3)	134 (28.0)	3 (2.7)
Other[d]	31 (5.3)	29 (6.1)	2 (1.8)
Mental health problem (N, %)[e]			
Depression	313 (53.1)	309 (64.5)	4 (3.6)
Dysthymia	208 (35.3)	200 (41.8)	8 (7.3)
Anxiety disorder	208 (35.3)	201 (42.0)	7 (6.4)
Personality disorder	105 (17.8)	99 (20.7)	6 (5.5)
Trauma- or stressor-related disorder	79 (13.4)	77 (16.1)	2 (1.8)
Autism or ADHD	77 (13.1)	76 (15.9)	1 (0.9)
Eating disorder	65 (11.0)	62 (12.9)	3 (2.7)
Obsessive-dompulsive disorder	46 (7.8)	43 (9.0)	3 (2.7)
Substance use disorder	36 (6.1)	35 (7.3)	1 (0.9)

Table 8.1. Demographic and clinical characteristics of study sample. (continued)

	Total sample	Mental healthcare service users[a]	Members of the general population[a]
Schizophrenia/psychosis	21 (3.6)	21 (4.4)	0 (0.0)
Other	19 (3.2)	16 (3.3)	3 (2.7)
Severity of current problems (N, %)[f]			
Mild	72 (12.2)	65 (13.6)	7 (6.4)
Moderate	241 (40.9)	232 (48.4)	9 (8.2)
Severe	139 (23.6)	137 (28.6)	2 (1.8)
Very severe	32 (5.4)	31 (6.5)	1 (0.9)
No problems anymore	15 (2.5)	14 (2.9)	1 (0.9)
No problems	90 (15.3)	0 (0.0)	90 (81.8)
MHQoL-7D			
Mean (SD)	12.3 (4.1)	11.5 (4.0)	15.5 (2.9)
Range	0 - 21	0 - 21	8-21
MHQoL-VAS			
Mean (SD)	6.0 (2.0)	5.7 (2.0)	7.5 (1.5)
Range	0 - 10	0 - 10	2 - 10

ADHD = Attention-Deficit/Hyperactivity Disorder; MHQoL = Mental Health Quality of Life questionnaire; SD = Standard Deviation; VAS = Visual Analogue Scale.

[a] Part of total sample.

[b] Lower, middle, and higher education refers to the ISCED [246] 2011 levels 0-2 (early childhood education, primary education, lower secondary education), 3-4 (upper secondary education, post-secondary non-tertiary education), and 5-8 (short-cycle tertiary education, bachelor or equivalent, master or equivalent, doctoral or equivalent), respectively.

[c] Some participants indicated that they visited more than one health professional for their mental health problems in the past 12 months.

[d] For example: community psychiatric nurse, hypnotherapist, vitality coach.

[e] Some participants indicated to have >1 mental health problem (mean number of mental health problems in total population was 2.4 (SD=1.4)).

[f] Severity was classified based on the own perception of participants.

Reliability

Table 8.2 presents the internal consistency reliability and test-retest reliability coefficients for the individual MHQoL-7D items. In the total sample, the Cronbach's alpha coefficient for the total MHQoL-7D was 0.85 and item-total correlations ranged from 0.48 to 0.71. None of the items could be deleted without a decrease of Cronbach's alpha. Test-retest reliability, as assessed by ICC, was 0.85 for the total MHQoL-7D. ICCs for individual items ranged from 0.51 to 0.77.

Convergent validity

Spearman's rank-order correlations between MHQoL scores and total scores of convergent measures are presented in Table 8.3. As hypothesised, the MHQoL showed

strong positive correlations with EQ-5D-5L, MANSA, and ICECAP-A scores. Moreover, there was a strong negative correlation between increasing MHQoL scores and psychopathology scores as measured by the BSI.

Table 8.2. Item-total correlations, alpha if item deleted and intraclass correlation coefficients for individual MHQoL-7D items.

Item	Total sample		Mental healthcare service users		Test-retest reliability subsample (N=195)
	Item-total correlation[a]	α if item deleted[a]	Item-total correlation[a]	α if item deleted[a]	ICC
1 - Self-image	0.69	0.81	0.67	0.79	0.73
2 - Independence	0.57	0.83	0.54	0.81	0.60
3 - Mood	0.69	0.81	0.65	0.79	0.70
4 - Relationships	0.49	0.84	0.44	0.83	0.70
5 - Daily activities	0.63	0.82	0.60	0.80	0.51
6 - Physical health	0.48	0.84	0.45	0.83	0.77
7 - Future	0.71	0.81	0.69	0.79	0.76

ICC = Intraclass Correlation Coefficient.
[a] Time point = baseline.
[b] All significant at P<0.001 (2-tailed).

Table 8.3. Spearman's rank-order correlations between MHQoL scores and total scores of convergent measures.[a,b]

Measure	Total sample	Mental healthcare service users	Members of the general population
EQ-5D-5L sum score	-0.58	-0.53	-0.47
EQ-5D-5L index	0.63	0.59	0.49
EQ-VAS	0.65	0.61	0.54
MANSA	0.75	0.71	0.69
ICECAP-A sum score	0.71	0.65	0.62
ICECAP-A index	0.71	0.66	0.62
BSI	-0.64	-0.57	-0.65

BSI = Brief Symptom Inventory; EQ-5D-5L = five-level EuroQol five-dimensional questionnaire; ICECAP-A = ICEpop CAPability measure for Adults; MANSA = Manchester Short Assessment of quality of life; VAS = Visual Analogue Scale.
[a] Time point = baseline.
[b] All significant at P<0.001 (2-tailed).

Known-group validity

A Mann-Whitney U test indicated that MHQoL-7D scores were significantly higher in participants with non-clinical psychopathology (Mdn=15) than in participants with clinical psychopathology (Mdn=11) (U=11,256; P<0.001; r=0.39). In addition, MHQoL-7D scores were significantly higher in participants with mild/moderate mental health

problems (Mdn=13) than in participants with severe mental health problems (Mdn=9) (U=12,300; P<0.001; r=0.42), and in members from the general population (Mdn=16) than in mental healthcare users (Mdn=12) (U=7.698; P<0.001; r=0.47).

Discussion

This paper reports on the development and psychometric evaluation of new quality of life measure for use in people with mental health problems, the Mental Health Quality of Life questionnaire (MHQoL). The MHQoL was designed to comprehensively provide information about the quality of life dimensions known to be relevant across and valued highly by people with mental health problems. Overall, the results of the present study suggest that the Dutch version of the MHQoL is a psychometrically sound measure of quality of life in Dutch people with mental health problems.

The face and content validity of the Dutch version of the MHQoL in Dutch people with mental health problems are supported by a multi-source, service user-oriented development process. Evaluation of the face and content validity by a focus group and online qualitative study confirmed the completeness and clarity of the MHQoL in this context. In addition, in the current study, the MHQoL demonstrated good internal consistency and good test-retest reliability over a one-week interval. Moreover, correlations between the Dutch version of the MHQoL and related measures supported convergent validity. As expected, higher scores on the MHQoL were strongly associated with higher scores on the ICECAP-A, EQ-5D-5L and MANSA. The MHQoL was more strongly associated with the ICECAP-A and MANSA than with the EQ-5D-5L. This is expected since the ICECAP-A and MANSA cover more dimensions included in the MHQoL compared to the EQ-5D-5L. In addition, there was a strong negative correlation between MHQoL scores and severity of mental health problems as measured by the BSI. Although quality of life has been found to be sensitive to variations in psychopathology [e.g. 243, 244], it is remarkable that the strength of the correlation between the MHQoL and BSI is comparable to the correlations between the MHQoL and other quality of life (EQ-5D-5L, MANSA) and well-being (ICECAP-A) measures. This finding raises the question what the differences between and interrelationships among quality of life, well-being and psychopathology are, also in terms of the underlying constructs. These is an interesting and important question, but one that falls beyond the scope of the current study and requires attention in future research. Known-group validity was supported by the ability of the MHQoL to detect significant differences in overall MHQoL levels between service users and the general population, between those reporting severe mental health problems and mild/moderate mental health

problems, and between those with clinical psychopathology and with non-clinical psychopathology.

The MHQoL offers several important advantages over most existing quality of life measures. The MHQoL was designed based on a comprehensive overview of the quality of life dimensions most relevant to people with mental health problems [40, 41]. Hence, the MHQoL is likely to be more sensitive to the benefits of mental healthcare interventions than generic quality of life measures. At the same time, it needs noting that this likely increase in sensitivity within the mental health domain may compromise the comparability of outcomes across sectors. However, in contrast to existing disease-specific quality of life measures, the MHQoL does still allow comparisons to be made across conditions *within* the mental health field. In addition, the MHQoL is relatively short and easy to complete by respondents in comparison to available quality of life measures used in people with mental health problems (average number of items=35 [237; Chapter 7]). The favourable ease of use of the MHQoL may support the use of the MHQoL in clinical and research settings alongside more clinically oriented measures, and would thereby accommodate the growing interest in measuring and monitoring the impact of mental health(care) on peoples' lives [233]. Although collecting 'traditional' outcomes, such as data on symptom remission, will remain essential, complementing it with outcome data on quality of life will offer a more complete understanding of the effectiveness of mental healthcare services, also from the perspective of those suffering from mental health problems. Moreover, the MHQoL can facilitate economic evaluations of mental health services, as further highlighted below.

The growing interest in comprehensive and sensitive outcome measures that can be used broadly in the mental health domain, may be underscored by the fact that recently more measures than only the MHQoL have been developed and introduced. To our knowledge, the only published examples of recently developed quality of life measures that cover all dimensions valued highly by people with mental health problems are the Recovering Quality of Life (ReQoL) measures [228]. Although the MHQoL and the ReQoL measures share the same goal, target population and theoretical basis (i.e. dimensions), they differ in a number of important ways, including the operationalization of their dimensions, the number of items (7 (MHQoL) vs. 10 (ReQoL-10) and 20 (ReQoL-20)), the recall period ("Today" (MHQoL) vs. "Last week" (ReQoL)), and the integration of the physical dimension in the measure (Integrated (MHQoL) vs. Supplemental (ReQoL)). The psychometric properties in terms of feasibility, reliability, validity and responsiveness of both ReQoL measures were reported to be satisfactory [228]. A direct comparison of the psychometric performance of the MHQoL and de ReQoL measures based on the published findings could not be performed because of

differences in sampling and measurement methods between the studies. Hence, we encourage future research to explore how the measures relate to one another and, for instance, which measure is preferred to be used in which context.

Several limitations to this study need to be acknowledged. First, as the presented study is a first psychometric evaluation of the MHQoL, future studies are needed to replicate and extend the findings from this initial evaluation. As the MHQoL was designed to adequately capture mental health-related quality of life and through that the benefits of mental healthcare interventions, in future studies special attention should be given to the evaluation of the sensitivity to change. In addition, future research is required to compare the sensitivity of the MHQoL to other (generic) quality of life measures and establish the effect of the use of a weighted sum score on the psychometric properties of the MHQoL. Second, the findings of the present study might have been subject to selection bias as participants were recruited by a market research company. Although people who voluntarily take part in online studies might differ from the general (patient) population, the sampling methodology resulted in a heterogeneous sample in terms of age, sex and education. Other consequences of the sampling procedure are that the rate of non-participation could not be determined, a relatively limited number of people with severe mental health problems participated, and a comprehensive psychiatric assessment by a mental health professional could not be performed, and hence, clinical and research diagnoses are missing. Future studies are needed to evaluate the psychometric properties in a clinically heterogeneous sample of mental healthcare service users. Third, in order to avoid subjective weighting of health states experienced over longer periods of time, and in line with other generic quality of life measures such as the EQ-5D, the recall period was set to "today". Recent research on issues related to different recall-periods and fluctuating health states indicates [247] that this choice may be influential and needs consideration also in the practical application of a measure. A main limitation of the here chosen recall period may be that fluctuations in quality of life may be missed and that obtained observations could be biased. This potential bias could, however, be reduced by administrating measures with a shorter recall period on a specific date, on a day with problems as well as on day without problems or by a more frequent administration of such measures [247]. In addition, this form of bias could be reduced by complementing the administration of the measure by diary completion in order to be able to assess whether the measure was administered on a day with or without problems. Fourth, in the present study, only the original Dutch version of the MHQoL was evaluated. English and German translations have been produced but are not yet tested for their psychometric properties. Broader validations of translated versions of the MHQoL in other countries are encouraged, in which cultural differences in relation to mental health should also be

considered. Fifth, as we tested the MHQoL in a sample of people aged 18 years and older, the MHQoL cannot be recommended for use in people younger than 18 without further psychometric evaluation, although, given the phrasing and domains, it may be considered potentially suitable for adolescents as well. Recommendations for future research include further psychometric testing, also in an international context, the development of a preference-based scoring algorithm to make the MHQoL suitable for use in cost-utility studies, and the direct comparison of the MHQoL with other recently developed quality of life measures for use in the mental health field such as the ReQoL measures. In addition, in order to increase the clinical relevance of the MHQoL, norm scores should be established to aid the interpretation of the MHQoL.

Notwithstanding these limitations, this study indicates that the MHQoL is a psychometrically sound measure in the Dutch context and, therefore, holds a promising capability as a simple, short and effective measure to assess quality of life in people with mental health problems. In order to make the MHQoL suitable for use in cost-utility analyses of mental healthcare interventions, preference weights will be estimated by use of a discrete choice experiment [248] in due course. By doing so, the MHQoL may facilitate sound economic evaluations of mental health interventions.

Appendix 8.1 – Mental Health Quality of Life questionnaire

Figure 8.1.1A. English translation of the MHQoL (part 1 of 2).

MHQoL

Please indicate below which statements best describe your situation **TODAY** by ticking **ONE** box in each of the seven subjects.

SELF-IMAGE

I think very positively about myself ☐
I think positively about myself ☐
I think negatively about myself ☐
I think very negatively about myself ☐

INDEPENDENCE *For example: freedom of choice, financial, co-decision making*

I am very satisfied with my level of independence ☐
I am satisfied with my level of independence ☐
I am dissatisfied with my level of independence ☐
I am very dissatisfied with my level of independence ☐

MOOD

I do not feel anxious, gloomy, or depressed ☐
I feel a little anxious, gloomy, or depressed ☐
I feel anxious, gloomy, or depressed ☐
I feel very anxious, gloomy, or depressed ☐

RELATIONSHIPS *For example: partner, children, family, friends*

I am very satisfied with my relationships ☐
I am satisfied with my relationships ☐
I am dissatisfied with my relationships ☐
I am very dissatisfied with my relationships ☐

DAILY ACTIVITIES *For example: work, study, household, leisure activities*

I am very satisfied with my daily activities ☐
I am satisfied with my daily activities ☐
I am dissatisfied with my daily activities ☐
I am very dissatisfied with my daily activities ☐

Figure 8.1.1B. English translation of the MHQoL (part 2 of 2).

PHYSICAL HEALTH

I have no physical health problems ☐
I have some physical health problems ☐
I have many physical health problems ☐
I have a great many physical health problems ☐

FUTURE

I am very optimistic about my future ☐
I am optimistic about my future ☐
I am gloomy about my future ☐
I am very gloomy about my future ☐

PSYCHOLOGICAL WELL-BEING

On the scale below, please indicate with an X how you rate your psychological well-being. 0 represents the worst imaginable psychological well-being, while 10 represents the best imaginable psychological well-being.

Worst imaginable psychological well-being — Best imaginable psychological well-being

10 General discussion

In light of the high prevalence and associated burden of mental health problems, national health systems strive to deliver accessible, efficient and high-quality mental healthcare. Since resources in healthcare are limited, they need to be used optimally, that is, in such a way that they yield most (health) benefits. Ensuring this is challenging and relates to both the organization of (mental) healthcare and the type of interventions offered in the mental healthcare system. A persistent challenge related to the former issue is the lengthy average treatment duration and pathway of patients with complex and severe mental health problems. In addition, questions have been raised on how to adequately identify, measure and value the outcomes of the interventions offered within the mental healthcare system. Adequate measurement of treatment outcomes could not only facilitate the optimization of tailored treatment pathways, but also a valid assessment of the effectiveness and cost-effectiveness of mental health interventions. This thesis addressed both issues by assessing whether instruments could be developed that have the potential to enhance the (evaluation of the) efficiency of mental healthcare.

This final Chapter discusses the main findings presented in this thesis in relation to the research questions outlined in Chapter 1. In addition, the strengths and limitations of this thesis are discussed and recommendations for further research and policy are provided.

Main findings

Chapters 2 and 3 examined *which individual patient indicators could facilitate the systematic and standardized early identification of patients in need of highly specialized mental healthcare* (research question 1). In Chapter 2, a systematic literature review was performed in order to identify pre-treatment patient characteristics that could facilitate the systematic and standardized early identification of patients with major depressive disorder in need of highly specialized care. The systematic literature review identified 48 characteristics of patients with major depressive disorder in need of highly specialized care that could be grouped into the following seven categories: depression severity, onset and (treatment) course, comorbid psychopathology, somatic comorbidity, childhood trauma, psychosocial functioning, and socio-demographics. Building on the results from Chapter 2, in Chapter 3, a concept mapping study was employed to appraise, refine, and complement the indicators derived from the systematic literature review with clinical expertise. In addition, the concept mapping study was performed to develop a consensus-based conceptual framework to inform the development of an instrument that could facilitate the systematic and standardized early identification

of patients with major depressive disorder in need of highly specialized mental healthcare. In total, 88 indicators of patients with major depressive disorder in need of highly specialized care were generated and categorized into the following ten conceptual domains: depression severity, onset and (treatment) course, comorbid personality disorder, comorbid substance use disorder, other psychiatric comorbidities, somatic comorbidity, maladaptive coping, childhood trauma, social factors, and psychosocial dysfunction. Building on the results from Chapters 2 and 3, **Chapters 4 and 5** assessed *how to develop and what the psychometric properties are of diagnosis-specific instruments that facilitate the systematic and standardized early identification of patients in need of highly specialized mental healthcare* (research question 2). In Chapter 4, an instrument aimed at facilitating clinicians in the systematic and standardized early identification of patients with major depressive disorder in need of highly specialized care was developed, the "Decision Tool Unipolar Depression" (DTUD). Each of the ten overarching domains that resulted from the concept mapping study (Chapter 3) was operationalized as a dichotomous item, resulting in the 10-item clinician-administered DTUD. In order to evaluate the psychometric properties of the DTUD, a cross-sectional multicenter study was conducted in a total of 243 patients with major depressive disorder. Overall, the DTUD demonstrated excellent feasibility, adequate inter-rater reliability, good convergent validity, and satisfactory criterion validity. Building on the theoretical foundations of, and insights from, the development of the DTUD, in Chapter 5, the Decision Tool Anxiety Disorders, OCD and PTSD (DTAOP) was developed and evaluated in terms of its psychometric properties. The DTAOP is an 8-item clinician-administered instrument that aims to aid clinicians in the systematic and standardized early identification of patients with an anxiety disorder, obsessive-compulsive disorder (OCD), or post-traumatic stress disorder (PTSD) in need of highly specialized mental healthcare. In line with the development process of the DTUD, a systematic literature review and a concept mapping study were carried out to inform the development of the DTAOP. In order to evaluate the psychometric properties of the DTAOP, a cross-sectional study in 454 patients with an anxiety disorder, OCD, or PTSD was carried out. The DTAOP demonstrated excellent feasibility, good convergent validity, satisfactory criterion validity, but weak inter-rater reliability. Based on the qualitative feedback provided by clinicians as part of the feasibility evaluation of the DTAOP, revisions and refinements of the wording and instructions were made in order to improve the item-level inter-rater reliability. The development and psychometric evaluation of the diagnosis-specific suggested that the allocation of patients to highly specialized mental healthcare settings, in general, may be guided by a core set of transdiagnostic patient factors. **Chapter 6,** therefore, examined *how to develop and what the psychometric properties are of an instrument that facilitates the systematic, early and standardized, transdiagnostic identification of patients in need of highly specialized mental healthcare*

(research question 3). Analysis of the overlapping criteria of the existing diagnosis-specific Decision Tools revealed five transdiagnostic criteria with which patients with a highly specialized care need could be detected (high severity level of the primary diagnosis, treatment-interfering psychiatric comorbidity, treatment-interfering somatic comorbidity, treatment-interfering psychosocial dysfunctioning, previous unsuccessful treatment of the current primary diagnosis in specialized mental healthcare). The criterion 'Severe or longstanding childhood trauma' was added to these initial five criteria due to its prognostic importance in patients with mental health problems. Consistent with the operationalization of the criteria of the diagnosis-specific Decision Tools, each of the transdiagnostic criteria was operationalized into a dichotomous scale item, resulting in the six-item Transdiagnostic Decision Tool. The Transdiagnostic Decision Tool was psychometrically evaluated in 505 patients with a somatic symptom disorder or PTSD. Overall, the Transdiagnostic Decision Tool demonstrated excellent feasibility, adequate inter-rater reliability, good convergent validity, and satisfactory criterion validity.

Despite the increasingly recognized importance of economic evaluations of mental healthcare interventions and the multitude of available quality of life instruments, concerns have been raised regarding the content validity of these instruments, and hence suitability for use in mental health. **Chapter 7** therefore evaluated *which quality of life instruments are currently used to measure the outcomes of mental healthcare interventions and what their content validity is* (research question 4). A systematic literature review was performed, which revealed a total of 44 quality of life instruments currently used in people with mental health problems. Based on previous work of Connell et al. [40, 41] an evaluation framework was established in order to assess the content validity of the identified instruments. The evaluation framework was used to assess whether the identified quality of life instruments cover the dimensions valued highly by people with mental health problems. The best coverage of the dimensions of the evaluation framework was by the WHOQOL-100, S-QoL 41, SQLS, EDQoL, QLI and IMHQOL, but none of the identified instruments fully covered the dimensions of the evaluation framework. The instruments with the best coverage of the dimensions of the evaluation framework lack a preference-based scoring algorithm at present. In light of the findings of Chapter 7, in **Chapter 8** it was examined *whether it was possible to develop a psychometrically sound mental health-related quality of life instrument* (research question 5). In order to inform the development of this instrument, a conceptual framework was constructed to serve as a theoretical basis. The conceptual framework was based on the work carried out by Connell et al. [40, 41] that provided a comprehensive overview of the quality of life dimensions most relevant to people with mental health problems. Informed by the items of the identified quality of life instruments used in people with

mental health problems (Chapter 7), scale items were generated for each of the dimensions of the conceptual framework. The resulting instrument was named the Mental Health Quality of Life questionnaire (MHQoL) and consists of a descriptive system, the MHQoL-7D, and a visual analog scale, the MHQoL-VAS. The MHQoL-7D comprises seven questions, covering seven dimensions (self-image, independence, mood, relationships, daily activities, physical health, future), each with four response levels (e.g., ranging from very satisfied to very dissatisfied). The MHQoL-VAS records the general psychological well-being on a horizontal scale ranging from zero ("worst imaginable psychological well-being") to ten ("best imaginable psychological well-being"). After its development, the MHQoL was evaluated for its psychometric properties in a heterogeneous population of 479 mental healthcare service users and 120 members from the general population. The MHQoL demonstrated to have favorable psychometric properties and showed promise as a simple and effective tool for assessing quality of life in people with mental health problems. In order to make the MHQoL suitable for use in cost-utility evaluations, the objective of **Chapter 9** was to evaluate whether *we could derive a valuation set with utility scores for a new mental health-related quality of life instrument* (research question 6). The valuation set was estimated using an efficient discrete choice experiment (DCE) with duration design that accommodated nonlinear time preferences. The DCE was embedded in a web-based survey and administered to a representative sample (N=1,505) of the Dutch adult population. In the obtained reference value set, utility values ranged from -0.687 for the worst state to 1 for the best state described with the MHQoL. The applied design enabled the generation of a preference-based value set that allows for the generation of an index value on a QALY scale anchored at 0 (death) and 1 (full health) and may, hence, be used in Dutch cost-utility analyses of mental healthcare interventions.

Strengths and limitations

The main strengths of this thesis lie in the relevance of the research question in an era of an increasing prevalence and burden of mental health problems and in the application of rigorous mixed-method approaches in the development and psychometric evaluation of the Decision Tools and the MHQoL. In the development and psychometric evaluation of the Decision Tools, we built on the theoretical foundations of, and insights from the applied methodology of earlier developed Decision Tools [173, 174]. By building on these insights, we were able to improve the previously applied methodology further. In addition, to derive a Dutch preference-based value set for the MHQoL we used a state-of-the-art discrete choice experiment (DCE). The application of this DCE allowed for the estimation of a preference-based value set that can be

used for the generation of an index value on a QALY scale anchored at 0 (death) and 1 (full health). As such, this thesis contributed to the wider application of such a state-of-the-art discrete choice experiments in the valuation of multi-attribute utility-based instruments. Despite these strengths, some limitations need to be considered when interpreting the findings and providing policy and research implications.

A first limitation concerns an important assumption that underlies the studies on the development of the Decision Tools presented in this thesis, namely that a quicker provision of appropriate care to patients with a highly specialized mental healthcare need enhances the clinical effectiveness and cost-effectiveness of mental healthcare. Although this assumption motivated the development of the Decision Tools, the benefit of matched care in patients with a highly specialized mental healthcare has not yet been studied. Use of the Decision Tools in daily clinical practice could, however, facilitate the early identification of patients in need of highly specialized treatment in order to facilitate a timely referral to appropriate treatment settings, which, in turn, is likely to improve the clinical and cost-effectiveness of care provided to these patients. Research in this area, aimed at confirming the benefits in terms of effectiveness and cost-effectiveness of matched care facilitated by Decision Tools, is strongly encouraged.

A second limitation concerns the Dutch context in which the Decision Tools were developed and psychometrically evaluated, which limits the generalizability of findings to other countries. For instance, the optimal cut-off value for each of the developed Decision Tools might differ across countries due to variations in the financing and organization of mental healthcare systems. In addition, the terminology used in the items and instructions of the items on the Decision Tools may be specific to the Dutch context. For instance, the terms "specialized mental healthcare" and "highly specialized mental healthcare" are used predominately in The Netherlands, although synonyms (e.g., secondary and tertiary mental healthcare) are available and used in other countries [25]. Hence, future studies are needed to assess the psychometric performance of the Decision Tools in other languages and countries in order to extend their cross-national robustness.

A third limitation concerns the relatively restricted field of application of the Decision Tools. The Decision Tools have been developed to aid clinicians in the early identification of patients with a highly specialized mental healthcare in need during the admission phase in specialized mental healthcare. Although the specific field of application for which the Decision Tools were developed facilitated a targeted development process and, in turn, may have enhanced the psychometric properties of the resulting Decision

Tools, it limits the application of the Decision Tools in other areas of interest. More specifically, broadening the field of application of the Decision Tools to primary care services may further enhance the provision of appropriate, timely care to patients with a highly specialized mental healthcare need. In addition, although the Decision Tools may facilitate a systematic and standardized early referral to highly specialized mental healthcare, they do not provide guidance in the specific type or intensity of treatment offered within the indicated level of care. The use of decision algorithms that guide the allocation of patients to specific treatments might further enhance the optimization of mental health(care). Hence, it is recommended to evaluate the possible use and benefits of Decision Tools across and within service levels in future studies.

A fourth limitation related to the studies on the development and psychometric evaluation of the Decision Tools presented in this thesis concerns the fact that the Decision Tools, to date, have been psychometrically evaluated only once. Even though the results of the psychometric properties of the diagnosis-specific and transdiagnostic Decision Tools presented in this thesis are promising, future studies are needed to replicate and extend these initial psychometric evaluations. More specifically, the psychometric evaluations of the Decision Tools that have been carried so far were cross-sectional in nature. Hence, psychometric features like test-retest reliability, predictive validity, and the sensitivity to treatment-related change have not yet been assessed and should be subject to future research. In addition, although the validity of the Transdiagnostic Decision Tool was found to approximate the validity of the available diagnosis-specific Decision Tools, a direct comparison of the Transdiagnostic Decision Tool and the diagnosis-specific tools has not yet been performed. Hence, future studies are required to directly compare the Transdiagnostic Decision Tool with the diagnosis-specific Decision Tools in terms of their psychometric performance in order to evaluate whether the Transdiagnostic Decision Tool may substitute the diagnosis-specific Decision Tools.

A fifth limitation of this thesis concerns the Dutch context in which the MHQoL was psychometrically evaluated and valued, limiting the cross-national generalizability of findings. Since the descriptive system of the MHQoL was based on a synthesis of international qualitative research on the dimensions important to the quality of life op people with mental health problems, the instrument itself may, after careful translation, be used in other countries. Although this may also suggest that the findings of the psychometric evaluation of the MHQoL could be generalizable to other countries, this needs confirmation in future studies. In addition, the Dutch tariff obtained for the MHQoL may differ between countries due to demographic and cultural differences

affecting preferences in the general public. Hence, further work is required to establish the MHQoL value sets in other countries.

A sixth and final limitation of this thesis the trade-off between comprehensiveness and relevance versus comparability of economic evaluations when using the MHQoL. The MHQoL was specifically designed as a comprehensive outcome measure for use in economic evaluations of mental healthcare interventions, as generic quality of life measures may not suffice in that context. This focus on mental health-related quality of life, also implies that the comparability of outcomes across sectors of economic evaluations is compromised. In other words, one cannot readily compare cost-per-QALY estimates from studies that used the MHQoL as an outcome measure with those that used a generic outcome measure like the EQ-5D, as they capture different concepts. Moreover, for decision makers, it is important to note that an ICER based on a "MHQoL QALY" may not necessarily need to be judged against the same monetary thresholds as ICERs based on generic QoL instruments. Hence, the benefit of being more appropriate, comprehensive and sensitive to the benefits of mental healthcare interventions comes at the price of reducing the comparability of outcomes across sectors and studies.

Implications for policy and future research

Notwithstanding the limitations of this thesis, it provided insight into whether instruments can be developed that have the potential to enhance the (evaluation of the) efficiency of mental healthcare. The findings of this thesis have several implications for policy and future research.

Given the demonstrated favorable psychometric properties of the developed Decision Tools, the developed Decision Tools may be recommended for use in daily clinical practice in the Netherlands. Note that, in fact, they are already being used in practice in several mental healthcare facilities across The Netherlands. These Decision Tools, therefore, appear to meet a need for accurate instruments that could facilitate the systematic and standardized early identification of patients in need of highly specialized mental healthcare. Although having demonstrated favourable psychometric properties, it is emphasized that the Decision Tools are intended to supplement rather than displace the "traditional" decision-making process when referring patients with mental health problems to highly specialized mental healthcare settings. In other words, the Decision Tools can provide indications of highly specialized care need, which, together with an assessment of the patient's individual circumstances, preferences and level of

motivation, could motivate a referral to a highly specialized mental healthcare setting. Despite their promising features, widespread implementation of instruments like the Decision Tools in daily clinical practice can be challenging [272]. As the criteria on the Decision Tools may seem to resemble the criteria many clinicians may implicitly take into account in the referral of patients to highly specialized mental healthcare settings, perhaps the biggest challenge in the implementation and use of Decision Tools is to increase confidence in instruments that aim to standardize procedures and assist clinicians' judgements in an objective fashion. One of the challenges for the mental healthcare field (like in other areas) may therefore be to stimulate a culture in which the acceptance of the use of instruments like the Decision Tools is strengthened. Recently, a first step has been taken to stimulate the implementation of the Decision Tools in daily clinical practice by including the Decision Tool Unipolar Depression (DTUD) in the Dutch standard of care for depressive disorders (in Dutch: Zorgstandard Depressieve stoornissen). In order to further stimulate the implementation of the Decision Tools in daily clinical practice, future research should be directed at potential barriers and facilitators of successful implementation of the Decision Tools. In addition, as mentioned in the previous paragraph, recommendations for future research related to the development and psychometric evaluation of the Decision Tools presented in this thesis concern the evaluation of their cross-national robustness, possible use and benefits of Decision Tools across and within service levels, and the way in which the Transdiagnostic Decision Tool may substitute the diagnosis-specific Decision Tools. Last but not least, an important area for future research is the evaluation of the benefit of matched care in patients with a highly specialized mental healthcare need. In this context, future research should be directed at the evaluation of whether Decision Tools indeed contribute to the shortening of the full treatment pathway of patients with complex and severe mental health problems to an adequate level of care, and to what extent this presupposed benefit of Decision Tools translates into the augmentation of the clinical and cost-effectiveness of care provided to these people.

The developed MHQoL was found to be a simple, short and psychometrically sound measure to assess quality of life in people with mental health problems. Ideally, this will be confirmed in future studies, in which the psychometric performance of the MHQoL is evaluated in a wide range of diagnoses and populations. Given its performance, the MHQoL can be recommended for use in the evaluation of the clinical effectiveness of mental healthcare interventions. Although the assessment of more commonly defined outcome domains in the mental health field, such as symptom remission, will remain essential and informative, complementing it with data on quality of life will offer a more complete understanding of the full spectrum of recovery of function and, hence, the effectiveness of mental healthcare interventions. Moreover, with the now avail-

able preference-based value set, the MHQoL facilitates cost-utility analyses of mental healthcare interventions. It should, however, be noted that in order to enhance the comparability of economic evaluations in healthcare, the EQ-5D is the recommended measure for use in cost-utility analyses in many countries worldwide [273, 274]. However, in patient populations in which the EQ-5D has been demonstrated to be insufficiently sensitive to condition-specific effects of interventions, some national health technology assessment agencies such as the Dutch National Health Care Institute recommend the use of alternative preference-based quality of life instruments in addition to the EQ-5D [273]. In that context and given the insights of this thesis, it is recommended to minimally include a broader, domain-specific preference-based quality of life instrument, such as the MHQoL, alongside the EQ-5D when performing cost-utility evaluations of mental healthcare interventions in order to make broader condition-specific effects of interventions visible and measurable. Although the use of such domain-specific, mental health-related quality of life instruments may compromise the comparability of outcomes across sectors, they do allow comparisons to be made across conditions within the mental health field and allow for the interpretation of the results in relation to generic quality of life measures such as the EQ-5D. In order to enhance the added value of preference-based, domain-specific outcomes, future research should be directed at the way in which and to what extent domain-specific outcomes (such as provided by the MHQoL) can be translated into future policy advice. In addition, recommendations for future research in this context specifically related to the MHQoL include the evaluation of its cross-national robustness, the estimation of preference-based value sets for other countries, and the comparison of the psychometric properties, including the sensitivity to change, of the MHQoL to other (generic) quality of life measures in the evaluation of mental healthcare interventions.

Final remarks

In light of the high prevalence and associated burden of mental health problems, national health systems strive to deliver accessible, efficient and high-quality mental healthcare. Since resources in healthcare are limited, they need to be used optimally, that is, in such a way that they yield most (health) benefits. This thesis contributed to this goal by developing instruments that have the potential to enhance the (evaluation of the) efficiency of mental healthcare. The findings indicate that psychometrically sound instruments can be developed for the early identification and adequate management of patients with mental health problems in need of highly specialized care. Their use in daily practice could enhance the systematic and standardized early identification of patients with a highly specialized mental healthcare need, and thereby

have the potential to enhance the efficiency og pathways of patients with complex and severe mental health problems to an adequate level of care. In addition, the developed MHQoL was found to hold a promising capability as a simple, short and psychometrically sound measure to assess quality of life in people with mental health problems. With the estimated preference-based value set, the MHQoL may now be used to evaluate whether patients in need of highly specialized care indeed benefit from quicker referrals and more tailored treatment pathways, but also to assess the effectiveness and cost-effectiveness of mental health interventions in general.

Concluding, this thesis contributed to the development, psychometric evaluation and valuation of instruments that have the potential to enhance the (evaluation of the) efficiency of mental healthcare. I hope that this thesis thereby contributed to the continuing challenge of the cost-effective use of scarce mental healthcare resources and ultimately to optimizing mental health(care).

References

1. James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, Abbastabar H, Abd-Allah F, Abdela J, Abdelalim A: 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. The Lancet 2018, 392(10159):1789-1858.
2. World Health Organization: The World Health Report 2001: Mental health: new understanding, new hope. 2001. Retrieved from [https://apps.who.int/iris/handle/10665/42390].
3. World Health Organization: Mental disorders - Key facts. 2019. Retrieved from [https://www.who.int/news-room/fact-sheets/detail/mental-disorders].
4. Soeteman DI, Verheul R, Busschbach JJ: The burden of disease in personality disorders: diagnosis-specific quality of life. J Personal Disord 2008, 22(3):259-268.
5. Hansson L: Quality of life in depression and anxiety. International Review of Psychiatry 2002, 14(3):185-189.
6. Wahlbeck K, Westman J, Nordentoft M, Gissler M, Laursen TM: Outcomes of Nordic mental health systems: life expectancy of patients with mental disorders. The British Journal of Psychiatry 2011, 199(6):453-458.
7. Bobinac A, Van Exel, N Job A, Rutten FF, Brouwer WB: Caring for and caring about: disentangling the caregiver effect and the family effect. J Health Econ 2010, 29(4):549-556.
8. Viana MC, Gruber MJ, Shahly V, Alhamzawi A, Alonso J, Andrade LH, Angermeyer MC, Benjet C, Bruffaerts R, Caldas-de-Almeida JM: Family burden related to mental and physical disorders in the world: results from the WHO World Mental Health (WMH) surveys. Brazilian Journal of Psychiatry 2013, 35:115-125.
9. Akbari M, Alavi M, Irajpour A, Maghsoudi J: Challenges of Family Caregivers of Patients with Mental Disorders in Iran: A Narrative Review. Iran J Nurs Midwifery Res 2018, 23(5):329-337.
10. Lee Y, Chatterton ML, Magnus A, Mohebbi M, Le LK, Mihalopoulos C: Cost of high prevalence mental disorders: Findings from the 2007 Australian National Survey of Mental Health and Wellbeing. Australian & New Zealand Journal of Psychiatry 2017, 51(12):1198-1211.
11. Oliva-Moreno J, López-Bastida J, Montejo-González AL, Osuna-Guerrero R, Duque-González B: The socioeconomic costs of mental illness in Spain. The European Journal of Health Economics 2009, 10(4):361-369.
12. Barrett B, Byford S, Seivewright H, Cooper S, Duggan C, Tyrer P: The assessment of dangerous and severe personality disorder: service use, cost, and consequences. The Journal of Forensic Psychiatry & Psychology 2009, 20(1):120-131.
13. Cary M, Butler S, Baruch G, Hickey N, Byford S: Economic evaluation of multisystemic therapy for young people at risk for continuing criminal activity in the UK. PloS one 2013, 8(4):e61070.
14. Kutcher S, McDougall A: Problems with access to adolescent mental health care can lead to dealings with the criminal justice system. Paediatrics & child health 2009, 14(1):15-18.
15. DeSocio J, Hootman J: Children's mental health and school success. The Journal of School Nursing 2004, 20(4):189-196.
16. Eisenberg D, Golberstein E, Hunt JB: Mental health and academic success in college. The BE Journal of Economic Analysis & Policy 2009, 9(1).
17. Day E, Shah R, Taylor RW, Marwood L, Nortey K, Harvey J, McAllister-Williams RH, Geddes JR, Barrera A, Young AH: A retrospective examination of care pathways in individuals with treatment-resistant depression. BJPsych Open 2021, 7(3).

18. Boenisch S, Kocalevent RD, Matschinger H, Mergl R, Wimmer-Brunauer C, Tauscher M, Kramer D, Hegerl U, Bramesfeld A: Who receives depression-specific treatment? A secondary data-based analysis of outpatient care received by over 780,000 statutory health-insured individuals diagnosed with depression. Soc Psychiatry Psychiatr Epidemiol 2012, 47(3):475-486.
19. Oquendo MA, Kamali M, Ellis SP, Grunebaum MF, Malone KM, Brodsky BS, Sackeim HA, Mann JJ: Adequacy of antidepressant treatment after discharge and the occurrence of suicidal acts in major depression: a prospective study. Am J Psychiatry 2002, 159(10):1746-1751.
20. Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ, Bruffaerts R, de Girolamo G, de Graaf R, Gureje O, Haro JM, Karam EG, Kessler RC, Kovess V, Lane MC, Lee S, Levinson D, Ono Y, Petukhova M, Posada-Villa J, Seedat S, Wells JE: Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. The Lancet 2007, 370(9590):841-850.
21. Kocsis JH, Gelenberg AJ, Rothbaum B, Klein DN, Trivedi MH, Manber R, Keller MB, Howland R, Thase ME: Chronic forms of major depression are still undertreated in the 21st century: systematic assessment of 801 patients presenting for treatment. J Affect Disord 2008, 110(1-2):55-61.
22. Brazier J: Measuring and valuing mental health for use in economic evaluation. J Health Serv Res Policy 2008, 13(3):70-75.
23. World Health Organization: Improving health systems and services for mental health. 2009. Retrieved from [https://apps.who.int/iris/handle/10665/44219].
24. Wasylenki D, Goering P, Cochrane J, Durbin J, Rogers J, Prendergast P: Tertiary mental health services: I. Key concepts. Can J Psychiatry 2000, 45(2):179-84.
25. Cochrane J, Goering P, Durbin J, Butterill D, Dumas J, Wasylenki D: Tertiary mental health services: II. Subpopulations and best practices for service delivery. Can J Psychiatry 2000, 45(2):185-190.
26. Saxena S, Thornicroft G, Knapp M, Whiteford H: Resources for mental health: scarcity, inequity, and inefficiency. The lancet 2007, 370(9590):878-889.
27. Bower P, Gilbody S: Stepped care in psychological therapies: access, effectiveness and efficiency: narrative literature review. The British Journal of Psychiatry 2005, 186(1):11-17.
28. Van Straten A, Tiemens B, Hakkaart L, Nolen W, Donker M: Stepped care vs. matched care for mood and anxiety disorders: a randomized trial in routine practice. Acta Psychiatr Scand 2006, 113(6):468-476.
29. Hirschfeld RM, Keller MB, Panico S, Arons BS, Barlow D, Davidoff F, Endicott J, Froom J, Goldstein M, Gorman JM: The National Depressive and Manic-Depressive Association consensus statement on the undertreatment of depression. JAMA 1997, 277(4):333-340.
30. Meyers BS, Sirey JA, Bruce M, Hamilton M, Raue P, Friedman SJ, Rickey C, Kakuma T, Carroll MK, Kiosses D: Predictors of early recovery from major depression among persons admitted to community-based clinics: an observational study. Arch Gen Psychiatry 2002, 59(8):729-735.
31. Lambert MJ, Whipple JL, Kleinstäuber M: Collecting and delivering progress feedback: A meta-analysis of routine outcome monitoring. Psychotherapy 2018, 55(4):520-537.
32. Slade M: What outcomes to measure in routine mental health services, and how to assess them: a systematic review. Australian & New Zealand Journal of Psychiatry 2002, 36(6):743-753.

33. Karimi M, Brazier J: Health, health-related quality of life, and quality of life: what is the difference? Pharmacoeconomics 2016, 34(7):645-649.
34. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW: *Methods for the economic evaluation of health care programmes:* Oxford university press; 2015.
35. Wisløff T, Hagen G, Hamidi V, Movik E, Klemp M, Olsen JA: Estimating QALY gains in applied studies: a review of cost-utility analyses published in 2010. Pharmacoeconomics 2014, 32(4):367-375.
36. Brazier J: Is the EQ-5D fit for purpose in mental health? The British Journal of Psychiatry 2010, 197(5):348-349.
37. Papaioannou D, Brazier J, Parry G: How valid and responsive are generic health status measures, such as EQ-5D and SF-36, in schizophrenia? A systematic review. Value in Health 2011, 14(6):907-920.
38. Saarni SI, Viertiö S, Perälä J, Koskinen S, Lönnqvist J, Suvisaari J: Quality of life of people with schizophrenia, bipolar disorder and other psychotic disorders. The British Journal of Psychiatry 2010, 197(5):386-394.
39. Brazier J, Connell J, Papaioannou D, Mukuria C, Mulhern B, Peasgood T, Jones ML, Paisley S, O'Cathain A, Barkham M, Knapp M, Byford S, Gilbody S, Parry G: A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. Health Technol Assess 2014, 18(34):1-188.
40. Connell J, Brazier J, O'Cathain A, Lloyd-Jones M, Paisley S: Quality of life of people with mental health problems: a synthesis of qualitative research. Health and quality of life outcomes 2012, 10(1):138.
41. Connell J, O'Cathain A, Brazier J: Measuring quality of life in mental health: Are we asking the right questions? Soc Sci Med 2014, 120:12-20.
42. Kessler RC, Aguilar-Gaxiola S, Alonso J, Chatterji S, Lee S, Ormel J, Ustun TB, Wang PS: The global burden of mental disorders: an update from the WHO World Mental Health (WMH) surveys. Epidemiol Psichiatr Soc 2009, 18(1):23-33.
43. Kessler RC, Angermeyer M, Anthony JC, DE Graaf R, Demyttenaere K, Gasquet I, DE Girolamo G, Gluzman S, Gureje O, Haro JM, Kawakami N, Karam A, Levinson D, Medina Mora ME, Oakley Browne MA, Posada-Villa J, Stein DJ, Adley Tsang CH, Aguilar-Gaxiola S, Alonso J, Lee S, Heeringa S, Pennell BE, Berglund P, Gruber MJ, Petukhova M, Chatterji S, Ustun TB: Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World Psychiatry 2007, 6(3):168-176.
44. Mathers CD, Loncar D: Projections of Global Mortality and Burden of Disease from 2002 to 2030. PLoS Med 2006, 3(11): doi:10.1371/journal.pmed.0030442.
45. Isacsson G, Boethius G, Bergman U: Low level of antidepressant prescription for people who later commit suicide: 15 years of experience from a population-based drug database in Sweden. Acta Psychiatr Scand 1992, 85(6):444-448.
46. Isometsa ET, Henriksson MM, Aro HM, Heikkinen ME, Kuoppasalmi KI, Lonnqvist JK: Suicide in major depression. Am J Psychiatry 1994, 151(4):530-536.
47. Suominen KH, Isometsa ET, Henriksson MM, Ostamo AI, Lonnqvist JK: Inadequate treatment for major depression both before and after attempted suicide. Am J Psychiatry 1998, 155(12):1778-1780.

48. Wang PS, Simon G, Kessler RC: The economic burden of depression and the cost-effectiveness of treatment. Int J Methods Psychiatr Res 2003, 12(1):22-33.
49. Meldon SW, Mion LC, Palmer RM, Drew BL, Connor JT, Lewicki LJ, Bass DM, Emerman CL: A Brief Risk-stratification Tool to Predict Repeat Emergency Department Visits and Hospitalizationsin Older Patients Discharged from the Emergency Department. Acad Emerg Med 2003, 10(3):224-232.
50. Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ: Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. Chest 2010, 137(2):263-272.
51. Cooperberg MR, Broering JM, Carroll PR: Risk assessment for prostate cancer metastasis and mortality at the time of diagnosis. J Natl Cancer Inst 2009, 101(12):878-887.
52. Garg S, Sarno G, Garcia-Garcia HM, Girasis C, Wykrzykowska J, Dawkins KD, Serruys PW, ARTS-II Investigators: A new tool for the risk stratification of patients with complex coronary artery disease: the Clinical SYNTAX Score. Circ Cardiovasc Interv 2010, 3(4):317-326.
53. Holzel L, Harter M, Reese C, Kriston L: Risk factors for chronic depression--a systematic review. J Affect Disord 2011, 129(1-3):1-13.
54. Riso LP, Miyatake RK, Thase ME: The search for determinants of chronic depression: a review of six factors. J Affect Disord 2002, 70(2):103-115.
55. Bennabi D, Aouizerate B, El-Hage W, Doumy O, Moliere F, Courtet P, Nieto I, Bellivier F, Bubrovsky M, Vaiva G, Holztmann J, Bougerol T, Richieri R, Lancon C, Camus V, Saba G, Haesbaert F, d'Amato T, Charpeaud T, Llorca PM, Leboyer M, Haffen E: Risk factors for treatment resistance in unipolar depression: A systematic review. J Affect Disord 2015, 171:137-141.
56. Lee AS: Better outcomes for depressive disorders? Psychol Med 2003, 33(05):769-774.
57. American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders (DSM-III)*. 1980.
58. American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders, -revised (DSM-III-R)*. 1987.
59. American Psychological Association: *Diagnostic and statistical manual of mental disorders (DSM-IV)*. 1994.
60. American Psychiatric Association: *Diagnostic and statistical manual of mental disorders DSM-IV-TR fourth edition (text revision)*. 2000.
61. American Psychiatric Association: *Diagnostic and statistical manual of mental disorders: DSM-5*. 2013.
62. World Health Organization: *International Classification of Diseases, Ninth Revision:* Geneva: World Health Organization; 1977.
63. World Health Organization: *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines:* Geneva: World Health Organization; 1992.
64. Spitzer RL, Endicott J, Robins E: Research diagnostic criteria: rationale and reliability. Arch Gen Psychiatry 1978, 35(6):773-782.
65. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D: The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009, 62(10).
66. VonVille HM. Organizing Search Data for Methods Reporting. 2014. Retrieved from [http://libguides.sph.uth. tmc.edu/excel_workbook_home/excel_SR_workbook].

67. National Heart, Blood and Lung Institute. Quality assessment tool for observational cohort and cross-sectional studies. 2014. Retrieved from [https://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools/case-control].
68. National Heart, Blood and Lung Institute. Quality Assessment Tool for Case-Control Studies. 2014. Retrieved from [http://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools/cohort].
69. Kaplan MJ, Klinetob NA: Childhood emotional trauma and chronic posttraumatic stress disorder in adult outpatients with treatment-resistant depression. J Nerv Ment Dis 2000, 188(9):596-601.
70. Souery D, Oswald P, Massat I, Bailer U, Bollen J, Demyttenaere K, Kasper S, Lecrubier Y, Montgomery S, Serretti A, Zohar J, Mendlewicz J, Group for the Study of Resistant Depression: Clinical factors associated with treatment resistance in major depressive disorder: results from a European multicenter study. J Clin Psychiatry 2007, 68(7):1062-1070.
71. Amital D, Fostick L, Silberman A, Beckman M, Spivak B: Serious life events among resistant and non-resistant MDD patients. J Affect Disord 2008, 110(3):260-264.
72. Dudek D, Rybakowski JK, Siwek M, Pawlowski T, Lojko D, Roczen R, Kiejna A: Risk factors of treatment resistance in major depression: Association with bipolarity. J Affect Disord 2010, 126(1-2):268-271.
73. Takahashi M, Shirayama Y, Muneoka K, Suzuki M, Sato K, Hashimoto K: Low openness on the Revised NEO Personality Inventory as a risk factor for treatment-resistant depression. PLoS ONE 2013, 8(9): e71964-7.
74. Takahashi M, Shirayama Y, Muneoka K, Suzuki M, Sato K, Hashimoto K: Personality traits as risk factors for treatment-resistant depression. PLoS One 2013, 8(5):e63756.
75. Riso LP, du Toit PL, Blandino JA, Penna S, Dacey S, Duin JS, Pacoe EM, Grant MM, Ulmer CS: Cognitive aspects of chronic depression. J Abnorm Psychol 2003, 112(1):72-80.
76. Gilmer WS, Trivedi MH, Rush AJ, Wisniewski SR, Luther J, Howland RH, Yohanna D, Khan A, Alpert J: Factors associated with chronic depressive episodes: a preliminary report from the STAR-D project. Acta Psychiatr Scand 2005, 112(6):425-433.
77. Wiersma JE, Hovens JG, van Oppen P, Giltay EJ, van Schaik DJ, Beekman AT, Penninx BW: The importance of childhood trauma and childhood life events for chronicity of depression in adults. J Clin Psychiatry 2009, 70(7):983-989.
78. Wiersma JE, van Oppen P, van Schaik DJ, van der Does, A J, Beekman AT, Penninx BW: Psychological characteristics of chronic depression: a longitudinal cohort study. J Clin Psychiatry 2011, 72(3):288-294.
79. Melartin TK, Rytsala HJ, Leskela US, Lestela-Mielonen PS, Sokero TP, Isometsa ET: Severity and comorbidity predict episode duration and recurrence of DSM-IV major depressive disorder. J Clin Psychiatry 2004, 65(6):810-819.
80. Solomon DA, Leon AC, Endicott J, Mueller TI, Coryell W, Shea MT, Keller MB: Psychosocial impairment and recurrence of major depression. Compr Psychiatry 2004, 45(6):423-430.
81. Bos EH, Bouhuys AL, Geerts E, Van Os TW, Van der Spoel, I D, Brouwer WH, Ormel J: Cognitive, physiological, and personality correlates of recurrence of depression. J Affect Disord 2005, 87(2-3):221-229.
82. Hollon SD, Shelton RC, Wisniewski S, Warden D, Biggs MM, Friedman ES, Husain M, Kupfer DJ, Nierenberg AA, Petersen TJ, Shores-Wilson K, Rush AJ: Presenting characteristics of depressed outpatients as a function of recurrence: preliminary findings from the STAR*D clinical trial. J Psychiatr Res 2006, 40(1):59-69.

83. Gerrits MM, van Oppen P, Leone SS, van Marwijk HW, van der Horst, H E, Penninx BW: Pain, not chronic disease, is associated with the recurrence of depressive and anxiety disorders. BMC Psychiatry 2014, 14:187-244.
84. Lamers F, Beekman AT, de Jonge P, Smit JH, Nolen WA, Penninx BW: One-year severity of depressive symptoms: results from the NESDA study. Psychiatry Res 2011, 190(2-3):226-231.
85. Butcher JN: Minnesota Multiphasic Personality Inventory (MMPI). 1989.
86. Endicott J, Nee J, Harrison W, Blumenthal R: Quality of Life Enjoyment and Satisfaction Questionnaire: a new measure. Psychopharmacol Bull. 1993.
87. Pollack MM, Alexander SR, Clarke N, Ruttimann UE, Tesselaar HM, Bachulis AC: Improved outcomes from tertiary center pediatric intensive care: a statewide comparison of tertiary and nontertiary care facilities. Crit Care Med 1991, 19(2):150-159.
88. Ryabchenko KA, Pepper CM, Jeglic EL, Griffith JW, Miller AB: Differences in course and comorbidity of recurrent depression in primary care and psychiatric populations. Depress Anxiety 2004, 20(3):153-154.
89. Hardeveld F, Spijker J, De Graaf R, Nolen W, Beekman A: Prevalence and predictors of recurrence of major depressive disorder in the adult population. Acta Psychiatr Scand 2010, 122(3):184-191.
90. van Weel-Baumgarten EM, Schers H, Van den Bosch W, Van den Hoogen H, Zitman F: Long-term follow-up of depression among patients in the community and in family practice settings: a systematic review. J Fam Pract 2000, 49(12):1113.
91. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS: Evidence based medicine: what it is and what it isn't. BMJ 1996, 312(7023):71-72.
92. Simon GE, Perlis RH: Personalized medicine for depression: can we match patients with treatments? Am J Psychiatry 2010, 167(12):1445-1455.
93. Meeuwissen JA, van der Feltz-Cornelis, Christina M, van Marwijk HW, Rijnders PB, Donker MC: A stepped care programme for depression management: an uncontrolled pre-post study in primary and secondary care in The Netherlands. International Journal of Integrated Care 2008, 8(1).
94. Von Korff M, Tiemens B: Individualized stepped care of chronic illness. West J Med 2000, 172(2):133-137.
95. Keller MB: Depression: a long-term illness. The British Journal of Psychiatry. 1994, 60:41-45.
96. Gask L, Khanna T: Ways of working at the interface between primary and specialist mental healthcare. Br J Psychiatry 2011, 198(1):3-5.
97. Ridgway N, Williams C: Cognitive behavioural therapy self-help for depression: an overview. Journal of Mental Health 2011, 20(6):593-603.
98. Delgadillo J, Moreea O, Lutz W: Different people respond differently to therapy: A demonstration using patient profiling and risk stratification. Behav Res Ther 2016, 79:15-22.
99. Dunlop BW, Mayberg HS: Neuroimaging-based biomarkers for treatment selection in major depressive disorder. Dialogues Clin Neurosci 2014, 16(4):479-490.
100. Lener MS, Iosifescu DV: In pursuit of neuroimaging biomarkers to guide treatment selection in major depressive disorder: a review of the literature. Ann N Y Acad Sci 2015, 1344(1):50-65.
101. Sotsky SM, Glass DR, Shea MT, Pilkonis PA, Collins JF, Elkin I, Watkins JT, Imber SD, Leber WR, Moyer J: Patient predictors of response to psychotherapy and pharmacotherapy: findings

in the NIMH Treatment of Depression Collaborative Research Program. Am J Psychiatry 1991, 148(8):997-1008.
102. McGrath CL, Kelley ME, Holtzheimer PE, Dunlop BW, Craighead WE, Franco AR, Craddock RC, Mayberg HS: Toward a neuroimaging treatment selection biomarker for major depressive disorder. JAMA psychiatry 2013, 70(8):821-829.
103. Evans KC, Dougherty DD, Pollack MH, Rauch SL: Using neuroimaging to predict treatment response in mood and anxiety disorders. Annals of Clinical Psychiatry 2006, 18(1):33-42.
104. van Krugten FC, Kaddouri M, Goorden M, van Balkom AJ, Bockting CL, Peeters FP, Hakkaart-van Roijen L, Decision Tool Unipolar Depression (DTUD) Consortium: Indicators of patients with major depressive disorder in need of highly specialized care: A systematic review. PLoS One 2017, 12(2):e0171659.
105. Trochim WM: An introduction to concept mapping for planning and evaluation. Eval Program Plann 1989, 12(1):1-16.
106. Johnsen JA, Biegel DE, Shafran R: Concept mapping in mental health: Uses and adaptations. Eval Program Plann 2000, 23(1):67-75.
107. Armstrong NP, Steffen JJ: The recovery promotion fidelity scale: assessing the organizational promotion of recovery. Community Ment Health J 2009, 45(3):163-170.
108. Conrad KJ, Iris M, Ridings JW, Rosen A, Fairman KP, Anetzberger GJ: Conceptual model and map of psychological abuse of older adults. J Elder Abuse Neglect 2011, 23(2):147-168.
109. Conrad KJ, Iris M, Ridings JW, Langley K, Anetzberger GJ: Self-report measure of psychological abuse of older adults. Gerontologist 2011, 51(3):354-366.
110. Corcoran K: The Oregon mental health referral checklists: Concept mapping the mental health needs of youth in the juvenile justice system. Brief Treatment and Crisis Intervention 2005, 5(1):9.
111. Kane M, Trochim WM: *Concept mapping for planning and evaluation:* Sage; 2007.
112. Trochim, W: The reliability of concept mapping. Annual conference of the American Evaluation Association. 1993.
113. Ward Jr JH: Hierarchical grouping to optimize an objective function. Journal of the American statistical association 1963, 58(301):236-244.
114. Shern DL, Trochim WM, LaComb CA: The use of concept mapping for assessing fidelity of model transfer: An example from psychiatric rehabilitation. Eval Program Plann 1995, 18(2):143-153.
115. van Krugten FC, Goorden M, van Balkom AJ, Spijker J, Brouwer WB, Hakkaart-van Roijen L: Indicators to facilitate the early identification of patients with major depressive disorder in need of highly specialized care: A concept mapping study. Depress Anxiety 2018, 35(4):346-352.
116. Hajjaj FM, Salek MS, Basra MK, Finlay AY: Non-clinical influences on clinical decision-making: a major challenge to evidence-based practice. J R Soc Med 2010, 103(5):178-187.
117. Rosas SR, Kane M: Quality and rigor of the concept mapping methodology: a pooled study analysis. Eval Program Plann 2012, 35(2):236-245.
118. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS: Evidence based medicine: what it is and what it isn't. BMJ 1996, 312(7023):71-72.
119. Ozomaro U, Wahlestedt C, Nemeroff CB: Personalized medicine in psychiatry: problems and promises. BMC medicine 2013, 11(1):132.
120. Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, Niederehe G, Thase ME, Lavori PW, Lebowitz BD: Acute and longer-term outcomes in depressed outpa-

tients requiring one or several treatment steps: a STAR* D report. Am J Psychiatry 2006, 163(11):1905-1917.
121. Keller MB: Depression: a long-term illness. The British Journal of Psychiatry 1994, 165(S26):9-15.
122. Haaga DA: Introduction to the special section on stepped care models in psychotherapy. J Consult Clin Psychol 2000, 68(4):547.
123. Richards DA, Bower P, Pagel C, Weaver A, Utley M, Cape J, Pilling S, Lovell K, Gilbody S, Leibowitz J: Delivering stepped care: an analysis of implementation in routine practice. Implementation Science 2012, 7(1):3.
124. Mathur S, Sutton J: Personalized medicine could transform healthcare. Biomedical reports 2017, 7(1):3-5.
125. Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH: The inventory of depressive symptomatology (IDS): psychometric properties. Psychol Med 1996, 26(3):477-486.
126. Svanborg P, Åsberg M: A comparison between the Beck Depression Inventory (BDI) and the self-rating version of the Montgomery Åsberg Depression Rating Scale (MADRS). J Affect Disord 2001, 64(2-3):203-216.
127. Williams JB: A structured interview guide for the Hamilton Depression Rating Scale. Arch Gen Psychiatry 1988, 45(8):742-747.
128. van Krugten FC, Goorden M, van Balkom AJ, van Oppen P, Ruhé HG, van Schaik DJ, Brouwer WB, Hakkaart-van Roijen L: The decision tool unipolar depression (DTUD): a new measure to facilitate the early identification of patients with major depressive disorder in need of highly specialized care. BMC Psychiatry 2019, 19(1):179.
129. Lobbestael J, Leurgans M, Arntz A: Inter-rater reliability of the Structured Clinical Interview for DSM-IV Axis I disorders (SCID I) and Axis II disorders (SCID II). Clinical psychology & psychotherapy 2011, 18(1):75-79.
130. Fekadu A, Wooderson S, Donaldson C, Markopoulou K, Masterson B, Poon L, Cleare AJ: A multidimensional tool to quantify treatment resistance in depression: the Maudsley staging method. J Clin Psychiatry 2009, 70(2):177-184.
131. Peeters FP, Ruhe HG, Wichers M, Abidi L, Kaub K, van der Lande, JH, Spijker J, Huibers MJ, Schene AH: The Dutch measure for quantification of treatment resistance in depression (DM-TRD): an extension of the Maudsley staging method. J Affect Disord 2016, 205:365-371.
132. Hayes AF, Krippendorff K: Answering the call for a standard reliability measure for coding data. Communication methods and measures 2007, 1(1):77-89.
133. Krippendorff K: *Content analysis: An introduction to its methodology:* Sage; 2012.
134. Lombard M, Snyder-Duch J, Bracken CC: Content analysis in mass communication: Assessment and reporting of intercoder reliability. Human communication research 2002, 28(4):587-604.
135. Cohen J: *Statistical power analysis for the behavioral sciences:* Academic press; 2013.
136. Youden WJ: Index for rating diagnostic tests. Cancer 1950, 3(1):32-35.
137. Barry MJ, VanSwearingen JM, Albright AL: Reliability and responsiveness of the Barry–Albright dystonia scale. Developmental Medicine & Child Neurology 1999, 41(6):404-411.
138. Støre-Valen J, Ryum T, Pedersen GA, Pripp AH, Jose PE, Karterud S: Does a web-based feedback training program result in improved reliability in clinicians' ratings of the Global Assessment of Functioning (GAF) Scale? Psychol Assess 2015, 27(3):865.

139. Fava M, A John Rush M, Alpert JE, Balasubramani G, Wisniewski SR, Carmin CN, Biggs MM, Zisook S, Leuchter A, Howland R: Difference in treatment outcome in outpatients with anxious versus nonanxious depression: a STAR* D report. Am J Psychiatry 2008, 165(3):342-351.
140. Frank E, Cassano GB, Rucci P, Thompson WK, Kraemer HC, Fagiolini A, Maggi L, Kupfer DJ, Shear MK, Houck PR: Predictors and moderators of time to remission of major depression with interpersonal psychotherapy and SSRI pharmacotherapy. Psychol Med 2011, 41(01):151-162.
141. Markowitz JC, Skodol AE, Petkova E, Cheng J, Sanislow CA, Grilo CM, Gunderson JG, McGlashan TH: Longitudinal effects of personality disorders on psychosocial functioning of patients with major depressive disorder. J Clin Psychiatry 2007, 68(2):186.
142. Penninx BW, Nolen WA, Lamers F, Zitman FG, Smit JH, Spinhoven P, Cuijpers P, de Jong PJ, van Marwijk HW, van der Meer K: Two-year course of depressive and anxiety disorders: results from the Netherlands Study of Depression and Anxiety (NESDA). J Affect Disord 2011, 133(1):76-85.
143. Newton-Howes G, Tyrer P, Johnson T, Mulder R, Kool S, Dekker J, Schoevers R: Influence of personality on the outcome of treatment in depression: systematic review and meta-analysis. J Personal Disord 2014, 28(4):577-593.
144. Brown GW, Harris TO, Kendrick T, Chatwin J, Craig TK, Kelly V, Mander H, Ring A, Wallace V, Uher R: Antidepressants, social adversity and outcome of depression in general practice. J Affect Disord 2010, 121(3):239-246.
145. Dezetter A, Briffault X, Bruffaerts R, De Graaf R, Alonso J, König H, Haro JM, De Girolamo G, Vilagut G, Kovess-Masféty V: Use of general practitioners versus mental health professionals in six European countries: the decisive role of the organization of mental health-care systems. Soc Psychiatry Psychiatr Epidemiol 2013, 48(1):137-149.
146. Dezetter A, Briffault X, Alonso J, Angermeyer MC, Bruffaerts R, de Girolamo G, De Graaf R, Haro JM, König HH, Kovess-Masfety V: Factors associated with use of psychiatrists and nonpsychiatrist providers by ESEMeD respondents in six European countries. Psychiatric Services 2011, 62(2):143-151.
147. Ayers CR, Sorrell JT, Thorp SR, Wetherell JL: Evidence-based psychological treatments for late-life anxiety. Psychol Aging 2007, 22(1):8.
148. Borkovec TD, Ruscio AM: Psychotherapy for generalized anxiety disorder. J Clin Psychiatry 2001, 62:37-45.
149. Roy-Byrne P, Craske MG, Sullivan G, Rose RD, Edlund MJ, Lang AJ, Bystritsky A, Welch SS, Chavira DA, Golinelli D: Delivery of evidence-based treatment for multiple anxiety disorders in primary care: a randomized controlled trial. JAMA 2010, 303(19):1921-1928.
150. Newman MG: Recommendations for a cost-offset model of psychotherapy allocation using generalized anxiety disorder as an example. J Consult Clin Psychol 2000, 68(4):549.
151. Baldwin DS, Waldman S, Allgulander C: Evidence-based pharmacological treatment of generalized anxiety disorder. International Journal of Neuropsychopharmacology 2011, 14(5):697-710.
152. Eisen JL, Sibrava NJ, Boisseau CL, Mancebo MC, Stout RL, Pinto A, Rasmussen SA: Five-year course of obsessive-compulsive disorder: predictors of remission and relapse. J Clin Psychiatry 2013, 74(3):233-239.
153. Taylor S, Abramowitz JS, McKay D: Non-adherence and non-response in the treatment of anxiety disorders. J Anxiety Disord 2012, 26(5):583-589.

154. Mughal AY, Devadas J, Ardman E, Levis B, Go VF, Gaynes BN: A systematic review of validated screening tools for anxiety disorders and PTSD in low to middle income countries. BMC Psychiatry 2020, 20(1):1-18.
155. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC: The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998, 59(20):22-33
156. Van Vliet I, De Beurs E: Het Mini Internationaal Neuropsychiatrisch Interview (MINI): Een kortgestructureerd diagnostisch psychiatrisch interview voor DSM-IV-en ICD-10-stoornissen. Tijdschrift voor psychiatrie. 2007, 49(6):393-397.
157. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X: Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of life research 2011, 20(10):1727-1736.
158. Versteegh MM, Vermeulen KM, Evers SM, de Wit GA, Prenger R, Stolk EA: Dutch Tariff for the Five-Level Version of EQ-5D. Value in Health 2016, 19(4):343-352.
159. Cramer V, Torgersen S, Kringlen E: Quality of life and anxiety disorders: a population study. J Nerv Ment Dis 2005, 193(3):196-202.
160. Chambless DL, Ollendick TH: Empirically supported psychological interventions: Controversies and evidence. Annu Rev Psychol 2001, 52(1):685-716.
161. Hollon SD, Ponniah K: A review of empirically supported psychological therapies for mood disorders in adults. Depress Anxiety 2010, 27(10):891-932.
162. Al-Janabi H, Flynn TN, Coast J: Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. Quality of Life Research 2012, 21(1):167-176.
163. Baker CB, Woods SW: Cost of treatment failure for major depression: direct costs of continued treatment. Adm Policy Ment Health Ment Health Serv Res 2001, 28(4):263–77.
164. Bloom DE, Cafiero E, Jané-Llopis E, Abrahams-Gessel S, Bloom LR, Fathima S, Feigl AB, Gaziano T, Hamandi A, Mowafi M: The global economic burden of noncommunicable diseases. 2012.
165. Diefenbach GJ, Tolin DF: The cost of illness associated with stepped care for obsessive-compulsive disorder. Journal of obsessive-compulsive and related disorders 2013, 2(2):144-148.
166. Hill JC, Whitehurst DG, Lewis M, Bryan S, Dunn KM, Foster NE, Konstantinou K, Main CJ, Mason E, Somerville S: Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. The Lancet 2011, 378(9802):1560-1571.
167. Goorden M, Huijbregts KM, van Marwijk HW, Beekman AT, van der Feltz-Cornelis, Christina M, Hakkaart-van Roijen L: Cost-utility of collaborative care for major depressive disorder in primary care in the Netherlands. J Psychosom Res 2015, 79(4):316-323.
168. Huijbregts KM, de Jong FJ, van Marwijk HW, Beekman AT, Adèr HJ, Hakkaart-van Roijen L, Unützer J, van der Feltz-Cornelis, Christina M: A target-driven collaborative care model for Major Depressive Disorder is effective in primary care in the Netherlands. A randomized clinical trial from the depression initiative. J Affect Disord 2013, 146(3):328-337.
169. Goorden M, Vlasveld MC, Anema JR, van Mechelen W, Beekman AT, Hoedeman R, van der Feltz-Cornelis, Christina M, Hakkaart-van Roijen L: Cost-utility analysis of a collaborative care intervention for major depressive disorder in an occupational healthcare setting. J Occup Rehabil 2014, 24(3):555-562.

170. Goorden M, van der Feltz-Cornelis, C M, van Steenbergen-Weijenburg KM, Horn EK, Beekman AT, Hakkaart-van Roijen L: Cost-utility of collaborative care for the treatment of comorbid major depressive disorder in outpatients with chronic physical conditions. A randomized controlled trial in the general hospital setting (CC-DIM). Neuropsychiatr Dis Treat 2017, 13:1881-1893.
171. Houtveen JH, van Broeckhuysen-Kloth S, Lintmeijer LL, Bühring ME, Geenen R: Intensive multidisciplinary treatment of severe somatoform disorder: a prospective evaluation. J Nerv Ment Dis 2015, 203(2):141-148.
172. de Vroege L, van Manen AS, Cees A, van der Feltz-Cornelis, Christina M: Complexity assessed by the INTERMED in patients with somatic symptom disorder visiting a specialized outpatient mental health care setting: a cross-sectional study. Psychosomatics 2017, 58(4):427-436.
173. Goorden M, Willemsen E, Bouwmans-Frijters C, Busschbach J, Noomx M, van der Feltz-Cornelis C, Uyl-de Groot C, Hakkaart-van Roijen L: Developing a decision tool to identify patients with personality disorders in need of highly specialized care. BMC Psychiatry 2017, 17(1):317.
174. Dingemans AE, Goorden M, Lötters FJ, Bouwmans C, Danner UN, van Elburg AA, van Furth EF, Hakkaart-van Roijen L: Development and validation of a decision tool for early identification of adult patients with severe and complex eating disorder psychopathology in need of highly specialized care. European Eating Disorders Review 2017, 25(5):366-372.
175. van Krugten FC, Kaddouri M, Goorden M, van Balkom AJ, Berretty EW, Cath DC, Hendriks G, Matthijssen SJ, Visser HA, van Vliet IM: Development and psychometric evaluation of the Decision Tool Anxiety Disorders, OCD and PTSD (DTAOP): Facilitating the early detection of patients in need of highly specialized care. Plos one 2021, 16(8):e0256384.
176. Flynn TN, Huynh E, Peters TJ, Al-Janabi H, Clemens S, Moody A, Coast J: Scoring the ICECAP-A capability instrument. Estimation of a UK general population tariff. Health Econ 2015, 24(3):258-269.
177. Dragatsi D, Norian I, Minkoff K: American Association of Community Psychiatrists Position Statement: Putting Patients First by Improving Treatment Planning and Reducing Administrative and Clinical Burden of Treatment Plan Documentation. Community Ment Health J 2019, 55(1):4-8.
178. Franken M, le Polain M, Cleemput I, Koopmanschap M: Similarities and differences between five European drug reimbursement systems. Int J Technol Assess Health Care 2012, 28(4):349-357.
179. Harris A, Buxton M, O'Brien B, Drummond, Rutten F, Drummond M: Using economic evidence in reimbursement decisions for health technologies: experience of 4 countries. Expert Review of Pharmacoeconomics & Outcomes Research 2001, 1(1): 7-12.
180. Weinstein MC, Torrance G, McGuire A: QALYs: the basics. Value in health 2009, 12:5-9.
181. Dolan P: Modeling valuations for EuroQol health states. Med Care 1997, 1095-1108.
182. Ware Jr JE, Sherbourne CD: The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. Med Care 1992:473-483.
183. Lamers L, Bouwmans C, van Straten A, Donker M, Hakkaart L: Comparison of EQ-5D and SF-6D utilities in mental health patients. Health Econ 2006, 15(11):1229-1236.
184. Mulhern B, Mukuria C, Barkham M, Knapp M, Byford S, Brazier J: Using generic preference-based measures in mental health: psychometric validity of the EQ-5D and SF-6D. The British Journal of Psychiatry 2014, 205(3):236-243.

185. Makai P, Brouwer WB, Koopmanschap MA, Stolk EA, Nieboer AP: Quality of life instruments for economic evaluations in health and social care for older people: a systematic review. Soc Sci Med 2014, 102:83-93.
186. Whoqol Group: Development of the World Health Organization WHOQOL-BREF quality of life assessment. Psychol Med 1998, 28(3):551-558.
187. Brazier J, Roberts J, Deverill M: The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002, 21(2):271-292.
188. Hawthorne G, Richardson J, Osborne R: The Assessment of Quality of Life (AQoL) instrument: a psychometric measure of health-related quality of life. Quality of Life Research 1999, 8(3):209-224.
189. Richardson J, Iezzi A, Khan MA, Maxwell A: Validity and reliability of the Assessment of Quality of Life (AQoL)-8D multi-attribute utility instrument. The Patient-Patient-Centered Outcomes Research 2014, 7(1):85-96.
190. Sintonen H: The 15D instrument of health-related quality of life: properties and applications. Ann Med 2001, 33(5):328-336.
191. Group TW: The World Health Organization quality of life assessment (WHOQOL): development and general psychometric properties. Soc Sci Med 1998, 46(12):1569-1585.
192. Auquier P, Simeoni M, Sapin C, Reine G, Aghababian V, Cramer J, Lancon C: Development and validation of a patient-based health-related quality of life questionnaire in schizophrenia: the S-QoL. Schizophr Res 2003, 63(1-2):137-149.
193. Wilkinson G, Hesdon B, Wild D, Cookson R, Farina C, Sharma V, Fitzpatrick R, Jenkinson C: Self-report quality of life measure for people with schizophrenia: the SQLS. The British Journal of Psychiatry 2000, 177(1):42-46.
194. Engel SG, Wittrock DA, Crosby RD, Wonderlich SA, Mitchell JE, Kolotkin RL: Development and psychometric validation of an eating disorder-specific health-related quality of life instrument. Int J Eat Disord 2006, 39(1):62-71.
195. Ferrans CE, Powers MJ: Quality of life index: development and psychometric properties. Advances in nursing science 1985, 8(1):15–24.
196. Long PW: Internet mental health quality of life scale (IMHQOL). Retrieved from [https://www.mentalhealth.com/qol/qol.html].
197. Ware Jr JE, Kosinski M, Keller SD: A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996, 34(3):220-233.
198. Schmidt S, Mühlan H, Power M: The EUROHIS-QOL 8-item index: psychometric results of a cross-cultural field study. The European Journal of Public Health 2005, 16(4):420-428.
199. Flanagan JC: A research approach to improving our quality of life. Am Psychol 1978, 33(2):138-147.
200. Flanagan JC: Measurement of quality of life: current state of the art. Arch Phys Med Rehabil 1982, 63(2):56-59.
201. Moriarty DG, Zack MM, Kobau R: The Centers for Disease Control and Prevention's Healthy Days Measures–Population tracking of perceived physical and mental health over time. Health and quality of life outcomes 2003, 1(1):37.
202. Cummins RA, Eckersley R, Pallant J, Van Vugt J, Misajon R: Developing a national index of subjective wellbeing: The Australian Unity Wellbeing Index. Soc Indicators Res 2003, 64(2):159-190.

203. Lindholt JS, Ventegodt S, Henneberg EW: Development and validation of QoL5 for clinical databases. A short, global and generic questionnaire based on an integrated theory of the quality of life. The European journal of surgery 2002, 168(2):107-113.
204. Lehman AF: A quality of life interview for the chronically mentally ill. Eval Program Plann 1988, 11(1):51-62.
205. 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.
206. Frisch MB: Quality-of-life-inventory. Encyclopedia of quality of life and well-being research 2014:5374-5377.
207. Oliver J, Huxley P, Priebe S, Kaiser W: Measuring the quality of life of severely mentally ill people using the Lancashire Quality of Life Profile. Soc Psychiatry Psychiatr Epidemiol 1997, 32(2):76-83.
208. Becker M, Diamond R, Sainfort F: A new patient focused index for measuring quality of life in persons with severe and persistent mental illness. Quality of life Research 1993, 2(4):239-251.
209. Baker F, Intagliata J: Quality of life in the evaluation of community support systems. Eval Program Plann 1982, 5(1):69-79.
210. Heinrichs DW, Hanlon TE, Carpenter Jr WT: The Quality of Life Scale: an instrument for rating the schizophrenic deficit syndrome. Schizophr Bull 1984, 10(3):388-398.
211. Boyer L, Simeoni M, Loundou A, D'Amato T, Reine G, Lancon C, Auquier P: The development of the S-QoL 18: a shortened quality of life questionnaire for patients with schizophrenia. Schizophr Res 2010, 121(1-3):241-250.
212. Brod M, Johnston J, Able S, Swindle R: Validation of the adult attention-deficit/hyperactivity disorder quality-of-life Scale (AAQoL): a disease-specific quality-of-life measure. Quality of life research 2006, 15(1):117-129.
213. Kazis LE, Lee A, Spiro A, Rogers W, Ren XS, Miller DR, Selim A, Hamed A, Haffer SC: Measurement comparisons of the medical outcomes study and veterans SF-36 health survey. Health Care Financ Rev 2004, 25(4):43-58.
214. Selim AJ, Rogers W, Fleishman JA, Qian SX, Fincke BG, Rothendler JA, Kazis LE: Updated US population standard for the Veterans RAND 12-item Health Survey (VR-12). Quality of Life Research 2009, 18(1):43-52.
215. Franz M, Fritz M, Gallhofer B, Meyer T: QLiS–development of a schizophrenia-specific quality-of-life scale. Health and quality of life outcomes 2012, 10(1):61.
216. Michalak EE, Murray G, CREST BD: Development of the QoL. BD: a disorder-specific scale to assess quality of life in bipolar disorder. Bipolar Disord 2010, 12(7):727-740.
217. Giner J, Ibáñez E, Baca E, Bobes J, Leal C, Cervera S: Desarrollo del cuestionario Sevilla de Calidad de Vida (CSCV). Actas Luso-Esp Neurol Psiquiatr 1997, 25(supl 2):11-23.
218. Vorstenbosch E, Bulten B, Bouman Y, Braun P: Forensic inpatient quality of life questionnaire. Nijmegen: Pompestichting; 2007.
219. Las Hayas C, Quintana JM, Padierna Á, Bilbao A, Muñoz P, Madrazo A, Urresti B, Cook EF: The new questionnaire Health-Related Quality of Life for Eating Disorders showed good validity and reliability. J Clin Epidemiol 2006, 59(2):192-200.
220. Las Hayas C, Quintana JM, Padierna JA, Bilbao A, Muñoz P: Use of rasch methodology to develop a short version of the Health Related Quality of life for Eating Disorders questionnaire: a prospective study. Health and Quality of life Outcomes 2010, 8(1):29.

221. Naber D, Karow A, Lambert M: Subjective well-being under neuroleptic treatment and its relevance for compliance. Acta Psychiatr Scand 2005, 111:29-34.
222. Power M, Quinn K, Schmidt S: Development of the WHOQOL-old module. Quality of life research 2005, 14(10):2197-2214.
223. Adair CE, Marcoux GC, Cram BS, Ewashen CJ, Chafe J, Cassin SE, Pinzon J, Gusella JL, Geller J, Scattolon Y: Development and multi-site validation of a new condition-specific quality of life measure for eating disorders. Health and Quality of Life Outcomes 2007, 5(1):23.
224. Cavanna AE, Schrag A, Morley D, Orth M, Robertson M, Joyce E, Critchley H, Selai C: The Gilles de la Tourette syndrome–quality of life scale (GTS-QOL): development and validation. Neurology 2008, 71(18):1410-1416.
225. Abraham SF, Brown T, Boyd C, Luscombe G, Russell J: Quality of life: eating disorders. Australian & New Zealand Journal of Psychiatry 2006, 40(2):150-155.
226. Touré M, Kouakou CR, Poder TG: Dimensions Used in Instruments for QALY Calculation: A Systematic Review. International Journal of Environmental Research and Public Health 2021, 18(9):4428.
227. Brazier JE, Yang Y, Tsuchiya A, Rowen DL: A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. The European journal of health economics 2010, 11(2):215-225.
228. Keetharuth AD, Brazier J, Connell J, Bjorner JB, Carlton J, Buck ET, Ricketts T, McKendrick K, Browne J, Croudace T: Recovering Quality of Life (ReQoL): a new generic self-reported outcome measure for use with people experiencing mental health difficulties. The British Journal of Psychiatry 2018, 212(1):42-49.
229. Keetharuth AD, Rowen D, Bjorner JB, Brazier J: Estimating a preference-based index for mental health from the recovering quality of life measure: Valuation of recovering quality of life utility index. Value in Health 2021, 24(2):281-290.
230. Mavranezouli I, Brazier JE, Young TA, Barkham M: Using Rasch analysis to form plausible health states amenable to valuation: the development of CORE-6D from a measure of common mental health problems (CORE-OM). Quality of Life Research 2011, 20(3):321-333.
231. Prinsen CA, Mokkink LB, Bouter LM, Alonso J, Patrick DL, De Vet HC, Terwee CB: COSMIN guideline for systematic reviews of patient-reported outcome measures. Quality of Life Research 2018, 27(5):1147-1157.
232. Fayers PM, Machin D: *Quality of life: The assessment, analysis and reporting of patient-reported outcomes:* John Wiley & Sons; 2015.
233. Shepherd G, Boardman J, Rinaldi M, Roberts G: Supporting recovery in mental health services: Quality and outcomes. London: Implementing Recovery Through Organisational Change; 2014.
234. Basu D: Quality-of-life issues in mental health care: Past, present, and future. German Journal of Psychiatry 2004, 7(3):35-43.
235. Fonagy P, Matthews R, Pilling S: Outcomes measures implementation best practice guidance. 2005.
236. Delespaul PA: Routine outcome measurement in the Netherlands–a focus on benchmarking. International Review of Psychiatry 2015, 27(4):320-328.
237. van Krugten FC, Feskens K, Busschbach JJ, Hakkaart-van Roijen L, Brouwer WB: Instruments to assess quality of life in people with mental health problems: a systematic review

and dimension analysis of generic, domain- and disease-specific instruments. Health and Quality of Life Outcomes. 2021, 19(249).
238. Keetharuth A, Brazier J, Connell J, Carlton J, Taylor Buck E, Ricketts T, Barkham M: Development and validation of the Recovering Quality of Life (ReQoL) outcome measures. In *EEPRU Technical Research Report 050. Policy Research Unit in Economic Evaluation of Health and Care Interventions.* University of Sheffield and York; 2017.
239. Derogatis LR, Spencer P: *Brief symptom inventory: BSI:* Pearson Upper Saddle River, NJ; 1993.
240. Derogatis LR, Unger R: Symptom checklist-90-revised. The Corsini encyclopedia of psychology 2010:1-2.
241. Peterson RA: A meta-analysis of Cronbach's coefficient alpha. Journal of consumer research 1994, 21(2):381-391.
242. Koo TK, Li MY: A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of chiropractic medicine 2016, 15(2):155-163.
243. da Silva Lima, Ana Flávia Barros, de Almeida Fleck, Marcelo Pio: Subsyndromal depression: an impact on quality of life? J Affect Disord 2007, 100(1-3):163-169.
244. Solanki RK, Singh P, Midha A, Chugh K: Schizophrenia: Impact on quality of life. Indian J Psychiatry 2008, 50(3):181-186.
245. Beurs Ed: Brief Symptom Inventory-BSI-, Brief Symptom Inventory 18,-BSI 18-, Handleiding herziene editie 2011. Leiden: PITS BV; 2011.
246. UNESCO Institute for Statistics: *International standard classification of education: ISCED 2011:* UNESCO Institute for Statistics Montreal; 2012.
247. Sanghera S, Coast J: Measuring Quality-Adjusted Life-Years When Health Fluctuates. Value in Health 2020, 23(3):343-350.
248. Bahrampour M, Byrnes J, Norman R, Scuffham PA, Downes M: Discrete choice experiments to generate utility values for multi-attribute utility instruments: a systematic review of methods. The European Journal of Health Economics 2020, 21(7):983-992.
249. Rose L, Mallinson RK, Walton-Moss B: A grounded theory of families responding to mental illness. West J Nurs Res 2002, 24(5):516-536.
250. Ekman M, Granstrom O, Omerov S, Jacob J, Landen M: The societal cost of schizophrenia in Sweden. J Ment Health Policy Econ 2013, 16(1):13-25.
251. Xu J, Wang J, Wimo A, Qiu C: The economic burden of mental disorders in China, 2005–2013: implications for health policy. BMC Psychiatry 2016, 16(1):1-9.
252. Lim KL, Jacobs P, Ohinmaa A, Schopflocher D, Dewa CS: A new population-based measure of the economic burden of mental illness in Canada. Chronic Dis Can 2008, 28(3):92-98.
253. Drummond M: Twenty years of using economic evaluations for drug reimbursement decisions: what has been achieved? J Health Polit Policy Law 2013, 38(6):1081-1102.
254. Edwards RT, Charles JM, Lloyd-Williams H: Public health economics: a systematic review of guidance for the economic evaluation of public health interventions and discussion of key methodological issues. BMC Public Health 2013, 13(1):1-13.
255. van Krugten F, Busschbach J, Versteegh M, Hakkaart-van Roijen L, Brouwer W: The Mental Health Quality of Life questionnaire (MHQoL): development and first psychometric evaluation of a new measure to assess quality of life in people with mental health problems. Quality of Life Research 2021:1-11.

256. Mulhern B, Norman R, Street DJ, Viney R: One method, many methodological choices: a structured review of discrete-choice experiments for health state valuation. Pharmacoeconomics 2019, 37(1):29-43.
257. Rowen D, Mulhern B, Stevens K, Vermaire JH: Estimating a Dutch value set for the pediatric preference-based CHU9D using a discrete choice experiment with duration. Value in Health 2018, 21(10):1234-1242.
258. Mulhern BJ, Bansback N, Norman R, Brazier J: Valuing the SF-6Dv2 classification system in the United Kingdom using a discrete-choice experiment with duration. Med Care 2020, 58(6):566-573.
259. Jonker MF, Donkers B, de Bekker-Grob EW, Stolk EA: Advocating a paradigm shift in health-state valuations: the estimation of time-preference corrected QALY tariffs. Value in Health 2018, 21(8):993-1001.
260. Jonker MF, Attema AE, Donkers B, Stolk EA, Versteegh MM: Are health state valuations from the general public biased? A test of health state reference dependency using self-assessed health and an efficient discrete choice experiment. Health Econ 2017, 26(12):1534-1547.
261. de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA: Sample size requirements for discrete-choice experiments in healthcare: a practical guide. The Patient-Patient-Centered Outcomes Research 2015, 8(5):373-384.
262. Viney R, Lancsar E, Louviere J: Discrete choice experiments to measure consumer preferences for health and healthcare. Expert review of pharmacoeconomics & outcomes research 2002, 2(4):319-326.
263. Jonker MF, Donkers B, de Bekker-Grob EW, Stolk EA: Effect of level overlap and color coding on attribute non-attendance in discrete choice experiments. Value in Health 2018, 21(7):767-771.
264. Jonker MF, Donkers B, de Bekker-Grob E, Stolk EA: Attribute level overlap (and color coding) can reduce task complexity, improve choice consistency, and decrease the dropout rate in discrete choice experiments. Health Econ 2019, 28(3):350-363.
265. Jonker M, Bliemer M: On the Optimization of Bayesian D-Efficient DCE Designs for the Estimation of QALY Tariffs That are Corrected for Nonlinear Time Preferences. Value in Health 2019, 22(10):1162-1169.
266. Mulhern B, Shah K, Janssen MFB, Longworth L, Ibbotson R: Valuing health using time trade-off and discrete choice experiment methods: does dimension order impact on health state values? Value in Health 2016, 19(2):210-217.
267. Mulhern B, Norman R, Lorgelly P, Lancsar E, Ratcliffe J, Brazier J, Viney R: Is dimension order important when valuing health states using discrete choice experiments including duration? Pharmacoeconomics 2017, 35(4):439-451.
268. Ericsson KA, Simon HA: How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. Mind, Culture, and Activity 1998, 5(3):178-186.
269. McFadden D: Conditional logit analysis of qualitative choice behavior. 1973.
270. Hakkaart-van Roijen L, Van der Linden N, Bouwmans C, Kanters T, Tan SS: Kostenhandleiding. Methodologie van kostenonderzoek en referentieprijzen voor economische evaluaties in de gezondheidszorg. 2015.
271. Cheung KL, Peter M, Smit C, de Vries H, Pieterse ME: The impact of non-response bias due to sampling in public health studies: a comparison of voluntary versus mandatory

recruitment in a Dutch national survey on adolescent health. BMC Public Health 2017, 17(1):1-10.

272. Levin SK, Nilsen P, Bendtsen P, Bulow P: Structured risk assessment instruments: a systematic review of implementation determinants. Psychiatry, Psychology and Law 2016, 23(4):602-628.
273. Nederland Z: Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg. Diemen: Zorginstituut Nederland; 2015.
274. National Institute for Clinical Excellence: Guide to the methods of technology appraisal. 2018. Retrieved from [http://www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalprocessguides/guidetothemethodsoftechnologyappraisal.jsp].

Summary

In light of the high prevalence and associated burden of mental health problems, national health systems strive to deliver accessible, efficient and high-quality mental healthcare. As resources in healthcare are limited, they need to be used in such a way that they yield most (health) benefits. Ensuring this is challenging and relates to both the organization of (mental) healthcare and the type of interventions offered in the mental healthcare system. A persistent challenge related to the former issue is the lengthy average treatment pathway of patients with complex and severe mental health problems. In addition, questions have been raised on how to adequately identify, measure and value the outcomes of the interventions offered within the mental healthcare system. Adequate measurement of treatment outcomes could not only facilitate the optimization of tailored treatment pathways, but also a valid assessment of the (cost-)effectiveness of mental health interventions. This thesis addressed both issues by developing instruments that have the potential to enhance the (evaluation of the) efficiency of mental healthcare.

Chapters 2 and 3 examined which individual patient indicators could facilitate the systematic and standardized early identification of patients in need of highly specialized mental healthcare. In Chapter 2, a systematic literature review was performed in order to identify pre-treatment patient indicators that could facilitate such an early identification of patients with major depressive disorder. The review identified 48 indicators of patients with major depressive disorder in need of highly specialized care. In Chapter 3, a concept mapping study was employed to appraise, refine, and complement the indicators derived from the systematic literature review with clinical expertise. In total, 88 indicators of patients with major depressive disorder in need of highly specialized care were generated and categorized into ten overarching conceptual domains.

Building on the results from Chapter 2 and 3, diagnosis-specific decision algorithms that could facilitate the systematic and standardized early identification of patients in need of highly specialized mental healthcare were developed and psychometrically evaluated in **Chapters 4 and 5**. In Chapter 4, a decision algorithm aimed at facilitating clinicians in the systematic and standardized early identification of patients with major depressive disorder in need of highly specialized care was developed, the "Decision Tool Unipolar Depression" (DTUD). Each of the ten domains that resulted from the concept mapping study (Chapter 3) was operationalized as a dichotomous item, resulting in the ten-item clinician-administered DTUD. In the psychometric evaluation, the DTUD demonstrated excellent feasibility, adequate inter-rater reliability, good convergent validity, and satisfactory criterion validity. In Chapter 5, a similar decision algorithm aimed at facilitating clinicians in the systematic and standardized early identification of patients with anxiety, obsessive compulsive disorder (OCD) and/or

post-traumatic stress disorder (PTSD) was developed and psychometrically evaluated, the "Decision Tool Anxiety Disorders, OCD and PTSD" (DTAOP). In line with the development process of the DTUD, a systematic literature review and a concept mapping study were carried out to inform the development of the descriptive system of the DTAOP. In the psychometric evaluation, the DTAOP demonstrated excellent feasibility, good convergent validity, satisfactory criterion validity, but weak inter-rater reliability. Based on the qualitative feedback provided by clinicians, revisions and refinements of the wording and instructions were made in order to improve the inter-rater reliability.

A comparison of the diagnosis-specific Decision Tools suggested that the allocation of patients to highly specialized mental healthcare settings could, potentially, be guided by a core set of transdiagnostic patient indicators. Therefore, in **Chapter 6**, a transdiagnostic Decision Tool was developed and tested for its psychometric properties. Based on the overlapping items of the developed diagnosis-specific Decision Tools and clinical expertise, a six-item Transdiagnostic Decision Tool was developed. In the psychometric evaluation, the Transdiagnostic Decision Tool demonstrated excellent feasibility, adequate inter-rater reliability, good convergent validity, and satisfactory criterion validity.

Despite the multitude of quality of life instruments used in people with mental health problems, concerns have been raised regarding the content validity of these instruments, and hence suitability for use in mental healthcare. **Chapter 7** therefore evaluated which quality of life instruments are currently used to measure the outcomes of mental healthcare interventions and what their content validity is. A systematic literature review was performed, which revealed a total of 44 quality of life instruments currently used in patients with mental health problems. Based on previous work of Connell et al. (2012) that provided a comprehensive overview of the quality of life dimensions most relevant to people with mental health problems, an evaluation framework was established in order to assess whether the identified quality of life instruments cover the dimensions valued highly by people with mental health problems. The WHOQOL-100, S-QoL 41, SQLS, EDQoL, QLI and IMHQOL best covered the dimensions of the evaluation framework; however, none of the identified instruments fully covered all dimensions that were deemed relevant.

In light of the findings of Chapter 7, a new quality of life measure for use in people with mental health problems was developed and psychometrically evaluated in **Chapter 8**, the "Mental Health Quality of Life questionnaire" (MHQoL). Based on previous work of Connell et al. (2012), a conceptual framework was constructed to serve as a theoretical basis for the measure. Scale items were generated for each of the dimensions

of the conceptual framework. The resulting MHQoL consists of a descriptive system comprising seven items that cover seven dimensions and a visual analog scale, the MHQoL-7D and MHQoL-VAS, respectively. In the psychometric evaluation, the MHQoL demonstrated favorable psychometric properties and, hence, showed promise as a feasible and effective tool for assessing quality of life in people with mental health problems.

In order to make the MHQoL suitable for use in economic evaluations of mental health interventions, a preference-based valuation set with utility scores for the measure was derived in **Chapter 9**. The valuation set was estimated using an efficient discrete choice experiment with duration design that was administered to a representative sample of the Dutch adult population. In the obtained value set, utility scores ranged from -0.687 for the worst state to 1 for the best state described by the MHQoL. The applied design enabled the generation of a value set that allows for the generation of an index value on a scale anchored at 0 (death) and 1 (full health), and hence can be used to assess the cost-effectiveness of mental health interventions.

This thesis contributed to the development, psychometric evaluation and valuation of instruments that have the potential to enhance the (evaluation of the) efficiency of mental healthcare. The results of this thesis indicated that psychometrically sound instruments can be developed for the early identification and adequate management of patients with mental health problems in need of highly specialized care. Their use in clinical routine could enhance the systematic and standardized early identification of patients with a highly specialized mental healthcare need, and thereby have the potential to enhance the tailoring of treatment pathways for patients with complex and severe mental health problems. In addition, the developed MHQoL was found to hold a promising capability as a feasible and psychometrically sound measure to assess quality of life in people with mental health problems. With the obtained preference-based value set, the MHQoL can be used to evaluate whether the quality of life of patients in need of highly specialized care indeed benefits from early referral and access to more tailored treatment pathways, but also to assess the (cost-)effectiveness of mental health interventions more broadly.

Samenvatting

Gezien de hoge prevalentie en de daarmee gepaard gaande last van psychische problemen, wordt binnen nationale zorgstelsels gestreefd naar het verlenen van toegankelijke, efficiënte en hoogwaardige geestelijke gezondheidszorg (GGZ). Aangezien de middelen in de zorg schaars zijn moeten deze zo worden ingezet dat ze zoveel mogelijk (gezondheids)winst opleveren. De bewerkstelliging hiervan is uitdagend en heeft zowel betrekking op de organisatie van de (geestelijke) gezondheidszorg als op het soort interventies dat in de GGZ wordt aangeboden. Een van de belangrijke uitdagingen gerelateerd aan de organisatie van de GGZ is het gemiddeld langdurige behandeltraject van patiënten met complexe en ernstige psychische problemen. Ook is het van belang dat de uitkomsten van de aangeboden interventies binnen de GGZ adequaat kunnen worden geïdentificeerd, gemeten en gewaardeerd. Adequate meting van behandelresultaten zou niet alleen de optimalisatie van behandeltrajecten kunnen vergemakkelijken, maar ook de valide beoordeling van de (kosten)effectiviteit van interventies in de GGZ kunnen bevorderen. In dit proefschrift zijn deze punten aan de orde gesteld door instrumenten te ontwikkelen die de potentie hebben om de (evaluatie van de) efficiëntie van de GGZ te bevorderen.

In **Hoofdstukken 2 en 3** is onderzocht welke patiëntkarakteristieken de systematische en gestandaardiseerde vroege identificatie van patiënten die hoogspecialistische GGZ nodig hebben zouden kunnen bevorderen. In Hoofdstuk 2 is een systematisch literatuuronderzoek uitgevoerd om patiëntkarakteristieken te identificeren die een dergelijke vroege identificatie van patiënten met een unipolaire depressie zouden kunnen vergemakkelijken. Door middel van het systematisch literatuuronderzoek werden 48 karakteristieken geïdentificeerd van patiënten met een unipolaire depressie met een behoefte aan hoogspecialistische GGZ. In Hoofdstuk 3 is een concept mapping studie uitgevoerd waarin de karakteristieken die werden gevonden in het systematisch literatuuronderzoek zijn gespecificeerd en aangevuld met klinische expertise. In totaal werden 88 unieke karakteristieken gevonden en gecategoriseerd in tien overkoepelende conceptuele domeinen.

Voortbouwend op de resultaten van Hoofdstukken 2 en 3 werden in **Hoofdstukken 4 en 5** diagnosespecifieke Decision Tools ontwikkeld en psychometrisch geëvalueerd. Deze Decision Tools zijn ontwikkeld met als doel de systematische en gestandaardiseerde vroege identificatie van patiënten met een behoefte aan hoogspecialistische GGZ te faciliteren. In Hoofdstuk 4 is een instrument ontwikkeld en psychometrisch geëvalueerd voor vroegtijdige herkenning van patiënten met een unipolaire depressie die behoefte hebben aan hoogspecialistische GGZ, de “Decision Tool Unipolaire Depressie” (DTUD). Elk van de tien conceptuele domeinen die resulteerden uit de concept mapping studie (Hoofdstuk 3) werd geoperationaliseerd als een dichotome

vraag. De psychometrische evaluatie toonde aan dat de DTUD een instrument is met een uitstekende hanteerbaarheid, adequate interbeoordelaarsbetrouwbaarheid, goede convergente validiteit en adequate criteriumvaliditeit. In Hoofdstuk 5 werd de "Decision Tool Angststoornis, Dwangstoornis en PTSS" (DTADP) ontwikkeld en psychometrisch geëvalueerd. De DTADP dient ter bevordering van de systematische en gestandaardiseerde vroege identificatie van patiënten met een angststoornis, dwangstoornis en/of posttraumatische stressstoornis (PTSS) die behoefte hebben aan hoogspecialistische GGZ. In lijn met het ontwikkelingsproces van de DTUD is voor de ontwikkeling van de DTADP een systematisch literatuuronderzoek en een concept mapping studie uitgevoerd. De psychometrische evaluatie toonde aan dat de DTADP een instrument is met een uitstekende hanteerbaarheid, goede convergente validiteit, adequate criteriumvaliditeit, maar zwakke interbeoordelaarsbetrouwbaarheid. Op basis van de kwalitatieve feedback van behandelaren gegeven tijdens de psychometrische evaluatie werden de vragen en bijbehorende instructies van de DTADP verder gespecificeerd ter bevordering van de interbeoordelaarsbetrouwbaarheid.

Een vergelijking van de ontwikkelde diagnosespecifieke Decision Tools suggereerde dat de vroegtijdige identificatie van patiënten met een behoefte aan hoogspecialistische GGZ mogelijk zou kunnen worden gefaciliteerd op basis van een kernset van transdiagnostische patiëntkarakteristieken. In **Hoofdstuk 6** is in dat licht een transdiagnostische Decision Tool ontwikkeld en psychometrisch geëvalueerd. Op basis van de overlappende vragen van de ontwikkelde diagnosespecifieke Decision Tools aangevuld met klinische expertise werd de uit zes vragen bestaande Transdiagnostische Decision Tool ontwikkeld. De psychometrische evaluatie toonde aan dat de Transdiagnostische Decision Tool een instrument is met een uitstekende hanteerbaarheid, adequate interbeoordelaarsbetrouwbaarheid, goede convergente validiteit en adequate criteriumvaliditeit.

Ondanks de veelheid aan kwaliteit van leven instrumenten die worden gebruikt bij mensen met psychische problemen, zijn er zorgen geuit over de inhoudsvaliditeit en daarmee de geschiktheid van deze instrumenten voor gebruik in de GGZ. In **Hoofdstuk 7** is om die reden geëvalueerd welke kwaliteit van leven instrumenten worden gebruikt om uitkomsten van GGZ-interventies te meten en wat de inhoudsvaliditeit van deze instrumenten is. Een systematisch literatuuronderzoek is uitgevoerd om kwaliteit van leven instrumenten te identificeren die worden gebruikt om de kwaliteit van leven te meten bij mensen met psychische problemen. In totaal werden 44 instrumenten gevonden. Gebaseerd op eerder werk van Connell et al. (2012) dat een overzicht geeft van de dimensies van kwaliteit van leven die het meest relevant zijn voor mensen met psychische problemen, werd een evaluatiekader opgesteld om de inhoudsvaliditeit van de 44 gevonden instrumenten te beoordelen. De WHOQOL-100, S-QoL 41, SQLS,

EDQoL, QLI en IMHQOL dekten het beste de dimensies van het evaluatiekader; geen van de geïdentificeerde instrumenten dekten echter volledig de relevant geachte dimensies.

In het licht van de bevindingen van Hoofdstuk 7 is in **Hoofdstuk 8** een nieuw kwaliteit van leven instrument ontwikkeld en psychometrisch geëvalueerd, de "Mental Health Quality of Life vragenlijst" (MHQoL). Op basis van eerder werk van Connell et al. (2012) werd een conceptueel raamwerk opgesteld dat diende als theoretische basis voor het te ontwikkelen instrument. Voor elk van de dimensies van het conceptuele raamwerk werd een vraag opgesteld. De MHQoL bestaat uit een beschrijvend gedeelte, de MHQoL-7D dat uit zeven vragen bestaat en een visueel analoge-schaal, de MHQoL-VAS. De psychometrische evaluatie toonde aan dat de MHQoL gunstige psychometrische eigenschappen heeft en daarmee een veelbelovend instrument is voor het meten van kwaliteit van leven bij mensen met psychische problemen.

Om de MHQoL geschikt te maken voor gebruik in economische evaluaties van interventies in de GGZ, is in **Hoofdstuk 9** een op voorkeuren gebaseerd tarief ontwikkeld door MHQoL-gezondheidstoestanden te laten waarderen door een representatieve steekproef van de Nederlandse bevolking. Het tarief werd geschat met behulp van een "discreet keuze-experiment". In het verkregen MHQoL-tarief varieerden de waarderingen, ook wel utiliteiten genoemd, van -0,687 voor de slechtste toestand tot 1 voor de beste toestand. Met het verkregen MHQoL-tarief kunnen utiliteiten worden gegenereerd op een schaal die loopt van 0 (dood) tot 1 (volledige gezondheid) en is de MHQoL geschikt voor gebruik in kosteneffectiviteitsanalyses van GGZ-interventies.

Dit proefschrift heeft een bijdrage geleverd aan de ontwikkeling, psychometrische evaluatie en waardering van instrumenten die de potentie hebben om de (evaluatie van de) efficiëntie van de GGZ te verbeteren. De resultaten van dit proefschrift illustreren dat psychometrisch goede instrumenten kunnen worden ontwikkeld voor de vroege identificatie van patiënten met een behoefte aan hoogspecialistische GGZ. Het gebruik van deze instrumenten zou de systematische en gestandaardiseerde vroege identificatie van patiënten met een behoefte aan hoogspecialistische GGZ kunnen verbeteren en daarmee kunnen bijdragen aan de optimalisatie van de behandeltrajecten van deze patiënten. Ook de ontwikkelde MHQoL blijkt een veelbelovend instrument voor het meten en waarderen van kwaliteit van leven bij mensen met psychische problemen. Met het verkregen tarief kan de MHQoL worden ingezet om te evalueren of patiënten met een behoefte aan hoogspecialistische GGZ inderdaad baat hebben bij meer gepersonaliseerde behandeltrajecten, maar ook om de (kosten-)effectiviteit van GGZ-interventies in het algemeen te beoordelen.

List of publications

van Krugten FC, Kaddouri M, Goorden M, van Balkom AJ, Bockting CL, Peeters FP, Hakkaart-van Roijen L: Indicators of patients with major depressive disorder in need of highly specialized care: A systematic review. PLoS One 2017, 12(2):e0171659.

van Krugten FC, Goorden M, van Balkom AJ, Spijker J, Brouwer WB, Hakkaart-van Roijen L: Indicators to facilitate the early identification of patients with major depressive disorder in need of highly specialized care: A concept mapping study. Depress Anxiety 2018, 35(4):346-352.

van Krugten FC, Goorden M, van Balkom AJ, van Oppen P, Ruhé HG, van Schaik DJ, Brouwer WB, Hakkaart-van Roijen L: The decision tool unipolar depression (DTUD): a new measure to facilitate the early identification of patients with major depressive disorder in need of highly specialized care. BMC Psychiatry 2019, 19(1):179.

van Krugten FC, Kaddouri M, Goorden M, van Balkom AJ, Berretty EW, Cath DC, Hendriks G, Matthijssen SJ, Visser HA, van Vliet IM, Brouwer WB, Hakkaart-van Roijen L: Development and psychometric evaluation of the Decision Tool Anxiety Disorders, OCD and PTSD (DTAOP): Facilitating the early detection of patients in need of highly specialized care. Plos one 2021, 16(8):e0256384.

van Krugten FC, van der Feltz-Cornelis CM, Boeschoten MA, van Broeckhuysen-Kloth SA, van Eck van der Sluijs JF, van Ee E, van Es SM, Schoorl M, Tak LM, Brouwer WB, Hakkaart-van Roijen L: Development and psychometric evaluation of the Transdiagnostic Decision Tool: matched care for patients with a mental disorder in need of highly specialised care. BJPsych Open 2020, 6(5).

van Krugten FC, Feskens K, Busschbach JJ, Hakkaart-van Roijen L, Brouwer WB: Instruments to assess quality of life in people with mental health problems: a systematic review and dimension analysis of generic, domain- and disease-specific instruments. Health and Quality of Life Outcomes. 2021, 19(249).

van Krugten FC, Busschbach JJ, Versteegh MM, Hakkaart-van Roijen L, Brouwer WB: The Mental Health Quality of Life questionnaire (MHQoL): development and first psychometric evaluation of a new measure to assess quality of life in people with mental health problems. Quality of Life Research 2021:1-11.

van Krugten FC, Himmler SF, Jonker MF, Hakkaart-van Roijen L, Brouwer WB: Estimating a preference-based value set for the Mental Health Quality of Life questionnaire (MHQoL). Submitted.

PhD Portfolio

PhD candidate:	Frédérique C.W. van Krugten
Dapartment:	Erasmus School of Health Policy & Management
PhD period:	2016-2021
Promotor:	Prof.dr. W.B.F. Brouwer
Copromotor:	Dr. L. Hakkaart-van Roijen

PhD Training

2014-2016	Research Master in Health Sciences, specialisation Health Economic Analysis *Netherlands Institute for Health Sciences, Erasmus Medical Center, Rotterdam, the Netherlands*
2016	Projectmanagement for PhD students *Erasmus Graduate School of Social Sciences and the Humanities, Erasmus University Rotterdam, the Netherlands*
2017	Diagnostic Research *Netherlands Institute for Health Sciences, Erasmus Medical Center, Rotterdam, the Netherlands*
2017	Basic Didactics *Risbo institute, Erasmus University Rotterdam, the Netherlands*
2017	Group dynamics and teaching *Risbo institute, Erasmus University Rotterdam, the Netherlands*
2017	Academic Writing for PhD Students *Erasmus Graduate School of Social Sciences and the Humanities, Erasmus University Rotterdam, the Netherlands*
2017	Psychiatric Epidemiology *Netherlands Institute for Health Sciences, Erasmus Medical Center, Rotterdam, the Netherlands*
2018	Self-presentation *Erasmus Graduate School of Social Sciences and the Humanities, Erasmus University Rotterdam, the Netherlands*
2019	Measurement of Patient Preferences using Discrete Choice Experiments *Erasmus School of Health Policy & Management, Erasmus University Rotterdam*

Teaching activities

2016-2018	Mentor First Year Mentor Program *Bachelor programme Health Sciences, Erasmus School of Health Policy & Management, Erasmus University Rotterdam, the Netherlands*
2016-2019	Instructor computer lab Health Technology Assessment *Master programme Health Economics, Policy & Law, Erasmus School of Health Policy & Management, Erasmus University Rotterdam, the Netherlands*
2016-2017	Supervisor Bachelor thesis *Bachelor programme Health Sciences: Health Policy & Management, Erasmus School of Health Policy & Management, Erasmus University Rotterdam, the Netherlands*

2017-2020	Coordinator Bachelor thesis *Bachelor programme Health Sciences: Health Policy & Management, Erasmus School of Health Policy & Management, Erasmus University Rotterdam, the Netherlands*
2018-2019	Coach Honours Class *Bachelor programme Health Sciences: Health Policy & Management, Erasmus School of Health Policy & Management, Erasmus University Rotterdam, the Netherlands*
2020-present	Coordinator Bachelor Graduation Project *Bachelor programme Health Sciences: Health Policy & Management, Erasmus School of Health Policy & Management, Erasmus University Rotterdam, the Netherlands*
2020-present	Supervisor Master thesis *Master programme Health Economics, Policy & Law, Erasmus School of Health Policy & Management, Erasmus University Rotterdam, the Netherlands*

Conference and symposium presentations

2016	Early Indicators of Patients with Major Depressive Disorder in Need of Highly Specialized Care - A Systematic Review (poster presentation) *International Society For Pharmacoeconomics and Outcomes Research 19th annual European Congress, Vienna, Austria*
2016	Feasibility, Reliability and Validity of the Decision Tool Unipolar Depression (DTUD) in Identifying Patients with Major Depressive Disorder in Need of Highly Specialized Care (poster presentation) *International Society For Pharmacoeconomics and Outcomes Research (ISPOR) 19th annual European Congress, Vienna, Austria*
2017	Decision Tools Ontwikkelen en Gebruiken (oral presentation) *Tenth anniversary TOPGGz Congress "Wissel-Werking: 10 jaar TOPGGz, zitten we op het goede spoor?", Amersfoort, the Netherlands*
2017	Feasibility, Reliability and Validity of the Decision Tool Unipolar Depression (DTUD) in Identifying Patients with Major Depressive Disorder in Need of Highly Specialized Care (oral presentation) *Thirteenth Workshop on Costs and Assessment in Psychiatry, Mental Health Policy and Economics, Venice, Italy*
2018	Indicators to Facilitate the Early Identification of Patients with Major Depressive Disorder in Need of Highly Specialized Care: A Concept Mapping Study (poster presentation) *26th European Congress of Psychiatry (EPA), Nice, France*
2018	Psychometric Evaluation of the Decision Tool Anxiety Disorders: Facilitating the Early Identification of Patients with an Anxiety Disorder in Need of Highly Specialized Care (poster presentation) *European Congress of Psychiatry (EPA), Nice, France*
2018	Decision Tools: Hulpmiddel ter voorkoming van onder- en overbehandeling (oral presentation) *Kennismiddag Transdiagnostisch Onderzoek Psychiatrie (TOP) consortium, Amersfoort, the Netherlands*
2019	Development and Psychometric Evaluation of the Transdiagnostic Decision Tool: Facilitating the Early Identification of Patients with Mental Health Problems in Need of Highly Specialized Care (oral presentation) *Fourteenth Workshop on Costs and Assessment in Psychiatry, Mental Health Policy and Economics, Venice, Italy*

2019	The Mental Health Quality of Life questionnaire (MHQoL): Development and First Psychometric Evaluation of a New Measure to Assess Quality of Life in People with Mental Health Problems (oral presentation) *International Health Economics Association (iHEA) World Congress, Basel, Switzerland*
2019	Meten van Kwaliteit van leven: de Mental Health Quality of Life questionnaire (MHQoL) (oral presentation) *Kennismiddag Transdiagnostisch Onderzoek Psychiatrie (TOP) consortium, Amersfoort, the Netherlands*

Other activities

2017-2019	Board member of the PhD Council young-ESHPM (yESHPM) *Erasmus School of Health Policy & Management, Erasmus University Rotterdam, the Netherlands*

About the author

Frédérique van Krugten (1991) obtained a Bachelor of Science degree (BSc) in Health Sciences in 2014 and a Master of Science degree (MSc) in Health Sciences (research, specialization health economics) in 2016 at the Erasmus University Rotterdam. She then became a PhD student at the Erasmus School of Health Policy and Management (ESHPM). Her research focuses on the development, psychometric evaluation and valuation of instruments to optimize mental healthcare. Her work has covered the development of various decision support algorithms for use in mental healthcare settings and the development, psychometric evaluation and valuation of the Mental Health Quality of Life questionnaire (MHQoL). The outcomes of her PhD research were published in peer-reviewed journals and presented at national and international conferences. During her PhD research, Frédérique was, as a lecturer and coordinator, involved in various bachelor and master courses of ESHPM. In addition, she served as a board member of the PhD association of ESHPM (YoungESHPM).