

From DEPARTMENT OF LEARNING, INFORMATICS, MEDICAL  
MANAGEMENT AND ETHICS  
Karolinska institutet, Stockholm, Sweden

**USE OF A STRUCTURED INTERVIEW TO SUPPORT  
DIAGNOSIS OF DEPRESSION AND ANXIETY DISORDERS  
IN PRIMARY CARE**

Agneta Pettersson



**Karolinska  
Institutet**

Stockholm 2017

All previously published papers were reproduced with permission from the publisher.

Published by Karolinska Institutet.

Cover photo: Agneta Pettersson

Printed by E-print

© Agneta Pettersson, 2017

ISBN 978-91-7676-632-3

# USE OF A STRUCTURED INTERVIEW TO SUPPORT DIAGNOSIS OF DEPRESSION AND ANXIETY DISORDERS IN PRIMARY CARE

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

**Agneta Pettersson**

*Principal Supervisor:*

Associate professor Ingvar Krakau  
Karolinska Institutet  
Department of Medicine  
Division of Clinical Epidemiology

*Opponent:*

Professor Per Nilsen  
Linköping University  
Department of Medical and Health Sciences  
Division of Community Medicine

*Co-supervisor(s):*

Professor Måns Rosén  
Karolinska Institutet  
Department of LIME  
Division of Medical Management Centre

*Examination Board:*

Professor Nils Lindefors  
Karolinska Institutet  
Department of Clinical Neurosciences  
Centre for Psychiatric Research

MD, Ph.D. Sonja Modin  
Karolinska Institutet  
Department of Neurobiology, Care Sciences and  
Society  
Division of Family Medicine and Primary Care

Professor Per Wändell  
Karolinska Institutet  
Department of Neurobiology, Care Sciences and  
Society  
Division of Family Medicine and Primary Care

Associate professor Rolf Wahlström  
Karolinska Institutet  
Department of Public Health Sciences  
Division of Global Health (IHCAR)

Associate professor Siw Carljford  
Linköping University  
Department of Medical and Health Sciences  
Division of Community Medicine



*Till min familj*

*“Bättre sent än aldrig!”*



# **ABSTRACT**

## **Aims**

The overall aims were to synthesize knowledge about which instruments for depression that are evidence based, and to develop a strategy, based on determinants of practice and theory, to support the use of these instruments in primary care.

## **Background**

A correct diagnosis is essential for appropriate management of depression and anxiety, the two most common mental disorders seen in primary care. However, the diagnosis can be difficult. Structured interviews and rating scales (instruments) have been proposed as a support in the consultation, but the evidence for their accuracy is unclear and their use in primary care limited.

## **Methods**

Determinants of practice were explored in a focus group study with  $n = 27$  family practitioners (FPs) in Västra Götaland. Systematic text condensation was used to analyse the discussions.

A systematic review was conducted in accordance to the PRISMA statement to determine which instruments that were evidence-based. Only studies with low or moderate risk for bias were included in the meta-analyses. The evidence for average sensitivity and specificity was rated with GRADE.

The acceptability of the evidence-based instrument, the Mini International Neuropsychiatric Interview (MINI), and of a supporting strategy was explored at three primary care centres (PCCs) in Stockholm. Data collection included a structured questionnaire to patients and interviewers ( $n = 125$  patients), semi structured interviews ( $n = 24$  patients and  $n = 3$  therapists) and focus groups ( $n = 17$  FPs). Qualitative content analysis and descriptive statistics were used in the data analysis. Findings across the participant groups were triangulated with the results from the questionnaires.

The main component of the intervention strategy was a task shift, where FPs could refer patients to a therapist for the MINI assessment and the results were fed back to the FP. The analyses were based on the interviews and focus groups, where  $n = 21$  patient interviews were relevant. Factors that influenced the referral process was identified with deductive content analysis guided by the COM-B model.

## Main Results

Several determinants of practice could influence when, and to what extent Swedish FPs use instruments for depression. These mainly concerned knowledge and attitudes of the individual FP. However, some actors outside primary care could influence the use.

The MINI had good diagnostic accuracy for depression. It was appreciated by FP, patients and therapists in primary care. The time for the assessment could be a problem for FPs who want to conduct the MINI.

The task shift was appropriate at one of the two PCCs. The inclination to refer was dependent on factors such as FPs' own knowledge and beliefs that referral gives benefits to the FP and the patient.

## Conclusions

The MINI can be a useful part of the consultation in primary care and is appreciated by the patients and interviewers. Referral to a therapist for MINI assessment can be feasible, depending on contextual factors.

Key words: depression, primary care, acceptability, determinants of practice, sensitivity and specificity, behaviour change, evidence based health care

## LIST OF SCIENTIFIC PAPERS

- I. Pettersson A, Bjorkelund C, Petersson EL. To score or not to score: a qualitative study on GPs views on the use of instruments for depression. *Fam Pract.* 2014; 31:215-221.
- II. Pettersson A, Bostrom KB, Gustavsson P, Ekselius L. Which instruments to support diagnosis of depression have sufficient accuracy? A systematic review. *Nord J Psychiatry.* 2015; 69:497-508.
- III. Pettersson A, Modin S, Wahlstrom R, af Winklerfelt Hammarberg S, Krakau I. The Mini-International Neuropsychiatric Interview is useful and well accepted as part of the clinical assessment for depression and anxiety in primary care: a mixed-methods study. Manuscript, submitted
- IV. Pettersson A, Modin S, Hasson H, Krakau I. Referral to a therapist for assessment of psychiatric problems – an interview study in primary care using COM-B. Manuscript

# TABLE OF CONTENTS

1	Background.....	9
	Primary care has become primary psychiatric care.....	9
	Depression and anxiety are common and risky but treatable.....	9
	Implementation and theory .....	19
	Implementation in practice.....	24
	Other theoretical underpinnings for the thesis.....	24
	Use of theoretical underpinnings in the thesis.....	26
2	Rationale and aims .....	27
3	Methods .....	29
	Design .....	29
	Setting .....	30
	The implementation object.....	31
	Implementation strategy.....	32
	Participants .....	33
	Data collection.....	34
	Data analyses .....	38
	Ethical considerations .....	41
4	Results.....	43
	Determinants for use of depression instruments in primary care .....	43
	The diagnostic accuracy for depression was sufficient for three instruments .....	44
	The MINI is useful and well accepted by patients and staff .....	45
	Referral to a therapist for MINI assessment can be feasible in primary care.....	46
5	Discussion.....	49
	Discussion of the results.....	49
	Methodological considerations .....	52
6	Conclusions .....	57
7	Clinical implications .....	58
8	Future research .....	59
9	Svensk sammanfattning .....	60
	Introduktion .....	60
	Syften .....	61
	Beskrivning av ingående artiklar och deras resultat.....	61
	Kommentarer .....	64
10	Acknowledgements .....	65
11	References .....	67

## LIST OF ABBREVIATIONS

BCW	Behaviour Change Wheel
CBT	Cognitive Behaviour Therapy
CFIR	Consolidated Framework for Implementation Research
COM-B	Competence, Opportunity, Motivation, Behaviour model
DSM-IV	Diagnostic and Statistical Manual for Mental Disorders, revision IV
EBM	Evidence based medicine
FP	Family practitioner
GRADE	The Grading of Recommendations Assessment Development and Evaluation
MINI	The Mini International Neuropsychiatric Interview
NICE	The National Institute for Health and Clinical Excellence
NPT	The Normalization Process Theory
PCC	Primary Care Center
PHQ-9	Patient Health Questionnaire, 9 items
PRIME-MD	Primary Care Evaluation of Mental Disorders
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
RCT	Randomized Controlled Trial
SBU	Swedish Council on Health Technology Assessments
SCID-I	Structured Clinical Interview for DSM-IV Axis I Disorders
STARD	Standards for Reporting of Diagnostic Accuracy Studies



# 1 BACKGROUND

The Background section describes the essential components for the thesis as follows:

Increasingly primary care has become primary psychiatric care (the setting); Diagnosis of depression and anxiety in primary care (the clinical problem); Use of a psychological instrument and barriers towards their use (a way to improve the diagnosis); and Models and frameworks.

## Primary care has become primary psychiatric care

Worldwide, increasing numbers of patients receive psychiatric treatment from FPs rather than from mental health specialists (for a review, see [1]). Some reasons are shortage of mental health care providers, insurance issues and patients' reluctance to visit psychiatric services [2]. The World Health Organization (WHO) and the World organization of Family Doctors (WONCA) support that mental health should be integrated into primary care as it offers advantages in terms of proximity, and person-centred and holistic services while minimizing stigma [3].

On the other hand, primary care has experienced problems in accommodating psychiatric care. The failure to identify co-morbidities that may interfere with treatment has been discussed [2]. Furthermore, disorders like depression and anxiety often follow a chronic course. Several strategies to manage patients with depression have been developed, e.g. collaborative care and use of the chronic care model [4]. A large survey study, conducted in the U.S., investigated to what extent four chronic conditions, depression, asthma, diabetes and congestive heart failure, were managed according to the principles of the chronic care model [5]. The study reported that PCCs used significantly less care management processes for depression compared to the somatic diseases. Disease registries were the most used part of the management, where 30 percent of the PCCs entered information [5].

## Depression and anxiety are common and risky but treatable

### Terminology and classification of depression and anxiety disorders

Mood and anxiety disorders are defined according to criteria in classification systems, most commonly the Diagnostic Statistical Manual for Mental Disorders (DSM, currently revision DSM-5) [6] and the International Statistical Classification of Diseases and Related Health Problems (ICD, currently revision ICD-10) [7]. Studies included in this thesis however, are based on the previous version of DSM, the DSM-IV [8].

The criteria are based on characteristic symptoms, their duration and severity. As an example, Box 1 shows the DSM-IV criteria for major depression.

Depressed mood and/or loss of interest or pleasure in life activities for at least 2 weeks, and at least five of the following symptoms that cause clinically significant impairment in social, work, or other important areas of functioning almost every day.

1. Depressed mood most of the day.
2. Diminished interest or pleasure in all or most activities (anhedonia).
3. Significant unintentional weight loss or gain.
4. Insomnia or sleeping too much.
5. Agitation or psychomotor retardation noticed by others.
6. Fatigue or loss of energy.
7. Feelings of worthlessness or excessive guilt.
8. Diminished ability to think or concentrate, or indecisiveness.
9. Recurrent thoughts of death

**Box 1** DSM-IV criteria for major depression (8).

Anxiety disorders in the DSM-IV included panic disorder, social phobia, specific phobias, generalized anxiety disorder (GAD), obsessive-compulsive disorder, (OCD), post-traumatic stress disorder (PTSD), and acute stress disorder [8].

The severity of depression and anxiety depends on the number of symptoms and how they affect daily function, e.g. relationship with the family and peers, and behaviour at work or school. However, many patients are in a grey zone. They experience substantial decrease in function and quality of life but do not fulfil sufficient criteria for a diagnosis. Therefore, definitions on subthreshold conditions have been created. Subthreshold depression [9] is defined as a depressive state, without the central DSM-IV criteria of depressed mood and anhedonia. The patients have at least two other symptoms, which have been present for more than two weeks, and are associated with significant social dysfunction. This is to be distinguished from subclinical conditions, in which individuals may have symptoms of a mental disorders without significant impairment or distress [10].

Depression and anxiety disorders often are chronic conditions. Penninx et. al. found that around 80 percent of patients with an episode of depression and 60 percent of patients with anxiety remitted [11]. The median time to remission was 6 months for a depressive disorder and 16 months for an anxiety disorder [11]. However, between 33 and 65 percent of primary care patients with depression relapse in long term studies [12, 13], and 10-17 percent follow a chronic course [13]. In primary care, around half of the patients with mental disorders have mild severity as reported in two studies [14, 15].

**Depression and anxiety disorders are common**

Several epidemiological studies have been conducted in Europe during the last 20 years. They have used different designs, aims and measurement methods [14-21]. A finding is that anxiety disorders are the most common mental disorders, with mood disorders as second [14]. Another finding is that the prevalence varies between countries, even when the same

measurement methods have been used [14, 17]. As an example, the WHO World Mental Health conducted a survey in 14 countries, including six European [14]. It found a 12-month prevalence for anxiety disorders between 5,8 and 12 percent in European countries, while mood disorders had a 12-month prevalence of 3,6 to 9,1 percent. Other studies show similar results [15-19, 21]. In contrast, a recent Swedish study reported far lower prevalence [20]. The study used data from a primary care database and found a 1,1 percent prevalence for mood disorders and 0,5 percent for anxiety disorders, measured as ICD-codes [20]. The authors cautioned that the figures may be an underestimation, since they are based on routine clinical examination.

### **Comorbidities are frequent**

Psychiatric comorbidity is the rule rather than the exception. The comorbidity between depression and anxiety in primary care populations has been investigated in several European studies [16, 18, 21, 22]. Anseau et. al. found that the prevalence of concomitant mood and anxiety disorders was 5,6 percent in a Belgian sample [16], and similar results have been reported from Sweden, Germany and Spain [15, 18, 21]. There was a strong association between an anxiety disorder and previous diagnoses of stress, sleep disorders and depression in a study conducted in the UK [22].

Furthermore, there are associations between severe somatic diseases and depression and anxiety [22-25], where e.g. depression is related to a higher frequency of diabetes complications [26].

### **Depression and anxiety disorders are risky**

As well documented in the literature, depression and anxiety disorders are associated with increased all causes mortality [27-30] as well as suicide [28, 31].

### **Depression and anxiety disorders are treatable**

Psychological and pharmacological therapies have proven effective for treatment of episodes of depression and anxiety disorders. The choice of treatment depends on the disorder, its severity and whether other treatments have been tried without success. Patients with mild disease might not need treatment, and watchful waiting or recommendations for life style are often recommended as an option during the first months [32].

Cognitive Behaviour Therapy, CBT, is indicated for depression as well as all anxiety disorders [32-37]. However, the content of the therapy is specific for each disorder, as well as the format (individual or group) and duration. Short psychodynamic therapies and applied relaxation are other psychological treatments that can be used for some of the disorders. Also for pharmacological treatment, the choice to some degree depends on the diagnosis. Accordingly, a correct diagnosis is vital for the treatment.

## Depression and anxiety are difficult to diagnose in the primary care setting

The dialogue between the FP and the patient is the corner stone for the diagnosis of patients with psychiatric symptoms. The patients are encouraged to tell their story in their own words. The FP may ask open ended questions for probing of symptoms and their duration. The FPs views on the diagnosis of depression and the diagnostic process has been the subject of two systematic reviews [38, 39]. According to these reviews, the FPs mostly saw depression as a reaction to life events. A minority considered depression as a reaction to a biochemical imbalance, which was also found in a Swedish study [40]. Professional qualities and communication skills were seen as essential and most FPs had established own routines and consultation styles. Intuition and “gut feeling” were helpful in the process.

This diagnostic process is associated with various problems. First, recognizing e.g. depression can be difficult since many depressive symptoms mirror symptoms of somatic diseases (See Table 1). Second, the information gathered could be influenced by factors related to e.g. patient and FP characteristics, the disorder with its severity and psychiatric comorbidities, and the communication [41-45].

Comorbid chronic somatic disease has been suggested to impact the recognition of depression. However, a systematic review by Menear et. al. found no evidence for a relationship between chronic somatic disease and likelihood of recognizing depression [46]. Rather, most studies suggested that no relationship exists [46]. On the other hand, symptoms such as not medically explained back pain, bodily weakness and permanent tiredness can contribute to the recognition of depression [47].

Shortage of time has been suggested as a reason for FPs’ failure to detect depression [48-50]. Pollock et. al. explored perspectives from FPs and patients on the consultation time [49, 50]. The nominal time for an appointment was 8-10 minutes, but the FPs overran the time if there was a clinical need [49]. The patients on the other hand, felt a pressure to keep the nominal time, which affected their communication negatively [50]. A systematic review concluded that there was some evidence that longer consultation times contributed to a more accurate diagnosis [48].

In summary, there are many factors that can aggravate the recognition of depression and anxiety disorders in primary care. Accordingly, the literature reports that FPs over diagnose as well as underdiagnose the disorders [51-55]. Mitchell et. al. conducted a systematic review, with 41 studies and more than 50 000 patients, where a clinical diagnosis of depression was compared to a psychiatric interview [53]. The FPs’ detection rate for depression was 50 percent, i.e. they missed half of the depressed patients. The FPs could correctly rule out depression in 80 percent of cases, i.e. a fifth would be falsely designated as depressed [53]. Another systematic review reported similar detection rates for anxiety disorders [55]. The literature suggests that severe depression is recognized to a higher extent [1, 56, 57].

## Proposed solutions for FPs to improve their diagnostic skills

Attempts have been made to help FPs improve their skills in diagnosing depression. One concept has been to provide training in depression or consultation techniques. A systematic review summarized the evidence for training of professionals working with mental health [58]. The systematic review included group, individual, and web-based training methods, mostly targeting family physicians. The review concluded that most methods had some positive effects on the skills. However, the impact on patient outcomes varied. Training knowledge and communication skills in groups of peers could improve mental health symptoms, while individual training had few long-lasting effects.

A second option is to provide standardized questionnaires based on DSM or ICD (designated “instruments” in the following text) to support the consultation [15, 59, 60]. Such instruments are the subject for this thesis.

### Does a diagnosis matter?

It seems reasonable that a correct diagnosis would imply a better management, and thus a better outcome for the patient. However, there are caveats which make the causal chain between correct diagnosis and patient outcome difficult. Patients may e.g. be reluctant to be treated, do not accept a mental disorder diagnosis or interrupt the treatment prematurely, or the first treatment option may not work. Such problems can obscure a positive effect of a correct diagnosis.

Many studies have concluded that recognition has transient or no effect on patient outcomes (See [56] for a review of older studies). Simon et. al. found that 36 percent of patients with major depression were recognized [56]. Recognized patients improved significantly more than non-recognized when measured after three months. However, after 12 months there were no differences [56]. A more recent Dutch study confirmed these findings [61].

More positive results were found by another Dutch study [62]. It investigated whether FPs recorded depression and anxiety diagnoses, and whether recording of the diagnosis influenced adherence to guidelines for depression and anxiety. Only 17,6 percent of patients with a depressive disorder and 8 percent of patients with an anxiety disorder had their diagnosis recorded in the electronic medical patient records. Appropriate treatment was given to most of the patients whose disorder was recognized while a minority of non-recognized patients were treated appropriately [62].

Instruments can potentially improve outcome for patients, but their diagnostic accuracy is unclear and FPs are reluctant to use them

There are two main groups of instruments available to support clinicians, disease-specific or including several psychiatric disorders. They are described in Table 1, which also lists some examples of each type of instruments.

**Table 1** Categories of instruments for detection and diagnosis of depression and anxiety disorders

<b>Instrument</b>	<b>Description</b>	<b>Duration</b>	<b>To be completed by</b>	<b>Examples</b>
Structured interview	Comprises modules for a range of diagnoses in the DSM- or ICD-systems. Each module contains items that together capture the criteria in the classifications. The patient can only respond yes or no.	Typically, 30 - 45 minutes	The interviewer, who can be a trained layman	MINI, PRIME-MD
Semi structured interview	Comprises modules for a range of diagnoses in the DSM- or ICD-systems. Each module contains items that together capture the criteria in the classifications. Questions are open-ended and give opportunities to expand on a subject.	1-2 hours	The interviewer; extensive training and medical background is needed	SCID-I
Case finding or screening instruments	Disease specific. Items cover criteria in the classification system	5- 10 minutes	Often the patient	HADS, PHQ-9, GAD-7
Assessment of severity	Disease specific. Items cover severity of problems and duration	5-10 minutes	Often the patient	MADRS, BDI-II,

MINI: Mini-International Neuropsychiatric Interview; PRIME-MD: The Primary Care Evaluation of Mental Disorders; SCID-I: Structured Clinical Interview for DSM-Axis I psychiatric disorders; HADS: Hospital Anxiety and Depression Scale; PHQ-9: Patient Health Questionnaire-9 item scale; GAD-7: General Anxiety Disorder -7 item scale; MADRS: Montgomery Asberg Depression Rating Scale; BDI-II: Beck Depression Inventory-II

The structured and semi structured interviews are based on the DSM-IV or ICD-10 classification systems. They comprise modules for a range of mental disorders, and each module contains items to capture all criteria in the classification system. Patients may fulfil one or several disorders in the interview. It should be noted that the included mental disorders to some extent vary between the interviews. As an example, the PRIME-MD

includes subthreshold depression and anxiety [63], which are not diagnoses according to the DSM-IV. In the remaining text, all interviews are labelled as “structured”.

Screening or case-finding instruments are disease specific. As for the interviews, they comprise items to capture the criteria of the disease. They can either be used to confirm a suspected diagnosis or to rule out a suspected diagnosis. Most of these instruments are to be completed by the patient (self-rated). A large number of instruments have been developed, with a content ranging from two items (e.g. the Whooley questions [64]) to comprehensive questionnaires such as the Beck Depression Inventory with 21 items [65]. The disease-specific instruments involve a rating from the patient to each of the included items. At the bottom line, these sum up to a score. Most often, developers of the various instruments have defined the threshold for disease, i.e. the minimum score.

A third category of instrument is aimed at assessing the severity of the disease. Such instruments are not a subject for this thesis.

### Expectations on instruments

Ultimately, the use of an instrument should result in a better outcome for the patient in need of treatment. The idea behind the use of structured interviews or questionnaires is to speed up the diagnostic process and facilitate the choice of an appropriate treatment. The benefit for the patient would be faster recovery or improvement.

Analyses and studies of the clinical effects of adding instruments to the consultation suffers from the same problems as for the association between a correct diagnosis and outcome. Having said that, few studies have investigated whether the use of instruments in the consultation improves the outcome for the patients, and the results are contradictory. It should be noted that all studies investigate the effects of case-finding or screening instruments for depression. One systematic review [66] found, that for unselected patients in the waiting room, the FPs management did not change when the FPs were informed about results of the screening before the consultation. There was no effect on improvement of depression. A second systematic review [67] found that quality improvement programs, where screening was an integral part of the management, led to a reduction in depression scores. However, a recent study, where patients were randomized to screening with the PHQ-9 to guide management, concluded that screening had no impact on outcome after 12 weeks [68].

One single study has investigated the impact of structured clinical interviews [69]. It is of principal interest although it took place in county funded psychiatric outpatient clinics and not in primary care. After a psychiatric evaluation, the patients were randomized to a semi structured interview, SCID, or to no interview. Physicians who treated patients from the SCID-group were more likely to order further tests, change diagnoses and change type of medication.

In summary, the effects of adding an instrument to the consultation are still unclear. However, there are positive findings that their use can have a positive impact on choice of treatments and patient outcome [62, 69].

**An instrument should have a good diagnostic accuracy**

The next issue is which instrument to choose. The evidence for the instruments was unclear, and the reason behind the SBU report on diagnosis of depression [70].

Properties of the instruments can be evaluated according to the same principles as used for diagnostic methods for somatic diseases. The test should be able to discriminate between those with a diagnosis and those without – they should have a good diagnostic accuracy. In evaluations, the test under investigation (index test) is compared with a reference standard. The reference standard by default is defined to classify 100 percent correct. For mental disorders, the Longitudinal Experts All Data, LEAD [71] and the similar Best Estimate procedures [72], have been developed to function as reference standards. They are composites, aimed at weighing together all available data for a diagnosis. Thus, LEAD includes an unstructured clinical interview, results from a structured interview and information from e.g. tests to exclude somatic diseases. All data is reviewed by a panel of experts. However, LEAD and Best Estimate are cumbersome procedures, and more commonly semi structured or structured interviews are used as reference standards.

The comparison gives the numbers of true and false positives (i.e. diseased) and true and false negatives. From these numbers, two core measures, the sensitivity and the specificity, can be calculated (See Table 2). They are correlated to each other and, in general, a high sensitivity is associated with a low specificity and vice versa.

**Table 2** Definition of sensitivity and specificity

	<b>Disease, according to the reference standard</b>	<b>Not disease, according to the reference standard</b>
<b>Disease, according to the index test</b>	TP (true positive)	FP (false positive)
<b>Not disease, according to the index test</b>	FN (false negative)	TN (true negative)

$$sensitivity = \frac{TP}{TP + FN}; specificity = \frac{TN}{FP + TN}$$

The sensitivity and specificity will depend on the chosen threshold, i.e. the border between disease and non-disease. Plotting the sensitivity and specificity (or rather 1- specificity) for all thresholds results in a Receiver Operating Curve (ROC). The shape of the curve can

inform developers of a test on the optimal choice for threshold. Lowering the threshold increases the sensitivity whilst the specificity decreases.

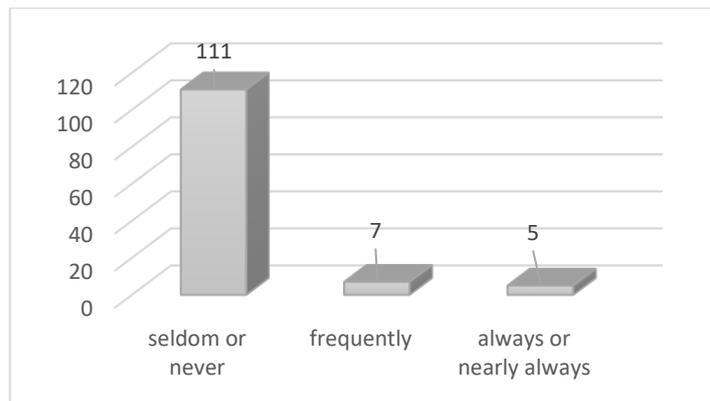
The optimal balance between sensitivity and specificity depends on the purpose of the test and a consideration of what gives more serious consequences for the patient: to falsely be labelled as having the disease or falsely labelled as not having the disease. For the case of anxiety and depression, a false positive diagnosis implies a treatment which is unnecessary and may cause side effects. On the other hand, a false negative diagnosis means that the patient will be deprived of a treatment and runs the risk of worsening.

#### Use of instruments in primary care

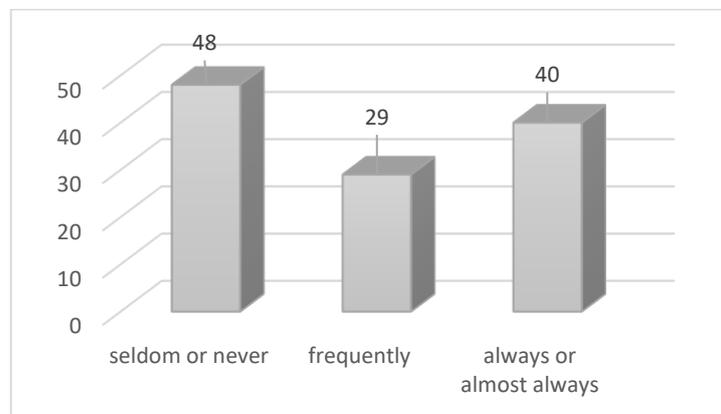
Internationally, use of instruments for mental disorders, primarily for depression, have received much attention from guideline developers as well as from the professions. The guidelines from the British National Institute for Health and Clinical Excellence, NICE, [33], and from the US agency ACHPR [73] recommended use of screening questionnaires. In the UK, screening questionnaires for patients with diabetes and ischaemic heart disease were part of the incentivized Quality and Outcomes Framework (QOF) [74].

In Sweden on the other hand, the National Board for Health and Social Welfare did not recommend the use of instruments in their National guidelines for anxiety and depression, published in 2010 [75]. The guidelines focused on a structured diagnostic process, based on diagnostic criteria [75]. However, two patient record studies, [76, 77] indicate that the use of diagnostic criteria is not the rule. In one of them, the prescription of anti-depressive drugs was usually based on a clinical interview without reference to diagnostic criteria [76]. The second report is based on records from one primary care centre [77]. Here, other ICD criteria than low mood seldom was recorded for patients who had been prescribed anti-depressive medication.

SBU conducted a national survey that was distributed to a representative sample of 300 FPs [70]. The response rate was 42 percent. Around 10 percent of the respondents used structured interviews regularly. With few exceptions, the PRIME-MD was selected (Figure 1). Around 60 percent of the respondents used screening questionnaires, usually HADS or MADRS-S (Figure 2).



**Figure 1** Frequency of use of structured interviews in a sample of Swedish FPs [70]. The y-axis depicts the number of respondents



**Figure 2** Frequency of use of screening questionnaires in a sample of Swedish FPs [70]. The y-axis depicts the number of respondents.

#### Attitudes and experiences of the use of instruments

As instruments became recommended, and even incentivized in some guidelines, studies began to investigate views on their use. A conclusion from these studies are that patients have more positive attitudes towards screening and severity instruments than the FPs [57, 78-84]. FPs, that had been thoroughly trained in the use of a screening instrument, were more positive though [85].

In the planning of the thesis, no studies on experiences from the use of structured interviews in primary care were identified.

## Implementation and theory

Broadly, implementation is defined as the process of putting a decision or plan into effect (Oxford Dictionary). Implementation research has been defined as “the scientific study of methods to promote the systematic uptake of clinical research findings and other evidence-based practices into routine practice, and hence to improve the quality (effectiveness, reliability, safety, appropriateness, equity, efficiency) of health care. It includes the study of influences on healthcare professional and organisational behaviour” [86]. Many similar definitions exist.

Interventions, such as education or feedback, have often been part of implementation efforts to support a change of behaviour. However, initially the interventions were chosen rather arbitrarily [87], and their effects were modest [88]. Furthermore, it was difficult to explain why implementation turned out to be a success or a failure.

Theoretical underpinnings for the interventions have been advocated in order to improve outcomes and understand mechanisms [87] - even if the evidence for their value in implementation research still is unclear [89]. Theoretical underpinnings can be theories, models and frameworks. Nilsen recently has categorized the various theoretical approaches in five groups [87]. These approaches and their use in implementation research are summarized in Table 3.

As the theoretical underpinnings can be used for several purposes, it is important to specify how theory should be applied [87, 89]. Furthermore, depending on the research question it may be necessary to use more than one theory or framework to capture all perspectives. On the other hand, handling two comprehensive frameworks may introduce redundancy and unnecessary work [89, 90].

### Determinants of practice

Determinants of practice are factors that may either hinder (barriers) or enhance (enablers or facilitators) implementation. The actual effect of a determinant will depend on factors such as the individuals targeted for the innovation, the intervention, and environmental factors. The combination of barriers and enablers will vary between settings [91] and may also change over time. A barrier during the first phases of implementation can turn into an enabler during later phases [92, 93]. Beforehand knowledge about barriers and enablers, and use of this knowledge to design the implementation strategy is believed to increase the likelihood of successful implementation [94].

### Determinant frameworks

Determinant frameworks comprise domains that have been shown to, or are assumed to, affect implementation outcome. Each domain contains several determinants. An early attempt for framework, by Greenhalgh et. al. [95], concluded that attributes of innovation, characteristics of the receiving organisation and the surrounding context, and the implementation process were important. Thereafter, a large number of frameworks have

**Table 3** Classification of theoretical underpinnings used in implementation science (adapted from Nilsen [87] with permission)

<b>Category</b>	<b>Description</b>	<b>Aim</b>
Process models	A model is a simplification of a phenomenon, closely related to theory. Process models specify stages in the process to translate research into practice	Descriptive. To guide the process
Determinant frameworks	A structure consisting of descriptive categories where the relations between them are presumed to account for a phenomenon	Descriptive. To specify determinants that acts as barriers and enablers and that influence implementation outcome
Evaluation frameworks	As determinant frameworks	Descriptive. To specify aspects of implementation that could be evaluated to determine success of implementation
Classic theories, originates from other research fields, e.g. psychology and sociology	A set of analytical principles. Made up of definitions of variables, a domain where the theory applies, a set of relationships between the variables and specific predictions	Explanatory and predictive (how and why specific relationships lead to specific events)
Implementation theories, developed by implementation researchers	As classic theories	As classic theories

been developed, with varying definitions, content and comprehensiveness [96]. Some frameworks are based on own experiences from implementation (e.g. [97]), while others have been constructed from theories, e.g. the Consolidated Framework for Implementation Research, CFIR [98] and the Theoretical Domains Framework (TDF) [99].

The Behaviour Change Wheel (BCW) is another synthesis of existing frameworks [100]. The wheel has three layers. The hub defines the determinants, based on a model, the COM-B. Thus, the COM-B identifies what needs to be changed to reach a desired behaviour. An

inner ring contains nine intervention functions. The intervention functions are education, persuasion, incentivisation, coercion, training, restriction, environmental restructuring, modelling and enablement. An outer ring identifies seven policy categories, communication, guidelines, fiscal measures, regulation, legislation, environmental planning and service provision. Intervention functions and policies should be chosen depending on the determinants.

In summary, most frameworks include the object to be implemented (the innovation), the target individual or organization, the end user, the context and the implementation strategy although they may be grouped and labelled differently.

### **The implementation object (the innovation)**

Implementation depends on the source of the object and its legitimacy. An example, where legitimacy played a role, is the implementation of the Swedish National guidelines for schizophrenia [101]. A study [102] found that the source, the National Board for Health and Welfare, had little legitimacy for the nurse end-users. The nurses considered the experts, that had written the guidelines, to be too far from the “floor”. The evidence for the implementation object is another, related, factor.

The object should have a good adaptability, i.e. it should be possible to modify it to meet local needs [95, 98]. Key concepts are the core components, which cannot be substituted or omitted, and the adaptable periphery which is possible to change.

### **Target individual or organization**

Depending on the research question, this domain can be interpreted as e.g. individuals or teams. The domain is not well developed in the CFIR [98], and a determinant such as motivation is not covered. The TDF [99] and the BCW [100] on the other hand, focus on the users of the object and include knowledge, emotions and skills.

### **Context and end-user**

Inner setting is a part of context in the CFIR. It describes the implementing organization, e.g. structural characteristics, communication within the organization, culture and implementation climate. Outer setting, in the CFIR, describes how the environment influences implementation. Care models, such as Shared decision-making and patient-centred care, imply that patients’ barriers and enablers to the object can have an impact. The inner and outer settings are captured in the Opportunity domain in the BCW framework.

### **The implementation strategy**

The purpose of the strategy is to facilitate implementation. The development of the strategy includes several steps. The problem must be clearly defined, e.g. improve detection of depression. A target behaviour that addresses the problem should be chosen, e.g. use of an

instrument. Determinants for the behaviour should be retrieved and interventions that are likely to address the determinants should be selected. Michie et. al. have developed the Behaviour Change Technique Taxonomy (v1) (BCTT (v1)) [103]. It compiles 93 behaviour change techniques (BCTs) in 16 groups. The BCTs can be linked to the intervention functions in the BCW [104]. As an example, for the intervention function Persuasion, the most frequently used BCTs are a credible source, information about consequences and feedback [104].

### *Tailoring the strategy*

The strategy should take determinants of practice into account, and be tailored to needs and opportunities for change [105]. Tailoring could be at the individual or more aggregated levels. Personalized tailoring is cumbersome, but could be successful in changing the behaviour. This is illustrated in some studies about improved management of depression in primary care [106-108].

Tailoring at group level, based on e.g. information from the literature, or from surveys and structured interviews with stakeholders, has also been investigated. A systematic review indicated that tailored interventions to support implementation of guidelines can change professional practice [109]. The impact on health care outcomes could not be estimated due to lack of studies [109]. Effect sizes generally ranged from small to moderate but some studies showed no effect. Furthermore, it was unclear how determinants and interventions should be selected [109]. More recently, results from the large Tailored Implementation in Chronic Diseases, TCID, project were published [110]. The interventions were based on a selection from a vast number of determinants [91], gathered from structured group meetings with stakeholders. Overall, TCID showed little observable impact on outcome and the authors raised concerns about tailoring as an approach to implementation [110].

### *Tailored strategies are often complex*

Most tailored implementation strategies are complex, as they encompass several interventions aiming at different determinants. The UK Medical Research Council guidance has developed a framework for evaluation of complex interventions [111]. It emphasizes the relations between implementation, mechanisms and context and the need to use theories. The MRC guidance recommends that the feasibility of a new intervention is investigated in pilot trials before proceeding to full scale implementation [111].

### **Theories and models**

Theories have partly been imported from other disciplines e.g. psychology and social work. The early Theory of Diffusion [112] has been influential in implementation science and has inspired frameworks as e.g. the CFIR [98]. The theory describes the innovation- decision process and highlights attributes of the innovation, adopter categories, communication networks and the importance of change agents [112].

Two examples of psychological behaviour change theories are the Theory of Planned Behaviour, TPB [113] and the Social Cognitive Theory [114]. TPB states that the behaviour is a result of an intention. The intention depends on the attitude towards the behaviour, beliefs about prevailing norms to the behaviour and a consideration whether the behaviour is possible [113]. The Social Cognitive Theory [114] defines that self-efficacy and outcome expectations affect individual behaviour, and that there is a continuous interaction between the individual and the environment.

Habit theory assumes that individuals that have developed a habit are less likely to change their behaviour [115]. If a behaviour is repeated in the same context, it gradually shifts from being guided by beliefs and attitudes to be triggered automatically. Behaviours that are repeated in the same context are difficult to change. Therefore, it is assumed that strategies that focus on changing the context have a greater probability of success [115]. One way of changing the context is a shift of professional roles. Systematic reviews have found that revision of professional roles could improve professional performance, while effects on patient outcomes were uncertain [116, 117].

Other theories have been developed within implementation science. The COM- B model (Capability, Opportunity, Motivation and Behaviour) has a starting point in motivation, defined as brain processes that direct behaviour [100]. Given sufficient motivation, two other components are required: individual capabilities (C), and opportunities (O). Opportunities cover factors that lie outside the individual. These three components and the behaviour interact with each other. The components can be further divided as shown in Box 2. The COM-B model is related to the more comprehensive TDF framework, and every domain in TDF can be linked to a component in COM-B [104].

<b>Component</b>	<b>Examples</b>
Capability, physical	Practical skill, e.g. taking a blood sample
Capability, psychological	Knowledge and skills, e.g. how to communicate with a patient
Opportunities, physical	e.g. allocation of time, staff resources
Opportunities, social	Interpersonal influences and norms
Motivation, reflective	Processes including intentions and beliefs about consequences
Motivation, automatic	Processes driven by emotional reactions and impulses

**Box 2** Components in the COM-B model

## Implementation in practice

The gap between current best practice, based on scientific evidence, and actual clinical care is well described in the literature. For health care as a whole, implementation of new evidence evolves slowly and is often only partial [118]. It has been shown that at least a third of patients do not receive care according to current best evidence, while a fifth receive care that is either not needed or even harmful [119]. Furthermore, only one third of the research evidence that informs guidelines is being adhered to [120].

## Implementation in primary care

Primary care organisations are complex. They vary, in e.g. composition of teams, organisational structure and working procedures, which can be a challenge for implementation [121]. Accordingly, many barriers to e.g. implementation of Evidence Based Medicine, EBM, and guidelines have been reported [122-126]. Lugtenberg et.al. identified barriers to a range of national guidelines, but the barriers varied between guidelines [123]. Toner et.al. found that less than 40 percent of FPs rated that NICE had at least moderate impact on their management of depression [125]. Common barriers regarding guidelines for anxiety and depression were lack of time [122, 125], no agreement on recommendations [122, 123] and patient factors [122, 125]. Gunn et.al. explored determinants of practice before implementing a best-practice depression care [126]. By use of the Normalization Process Theory [127], they found four determinants that would have an impact on the implementation: a lack of shared understanding between staff of what depression is, a lack of agreement about treatments, how the care should be organized and how patients should be monitored [126].

## Other theoretical underpinnings for the thesis

### Mixed methods

Mixed methods have been proclaimed a third paradigm [128] and much of the current literature falls back on a conceptual framework developed by Greene et.al. [129]. The field of mixed methods was formally established in the late 1980s. It has grown and expanded to various disciplines such as social work, education, psychology and medicine.

There is no agreed definition of mixed methods though, and what should be considered as mixed methods and what should not [130]. Most definitions state that mixed method research integrates quantitative and qualitative approaches to investigate a phenomenon - in the same study [130-132]. The underlying assumption is that mixed methods can address research questions more comprehensively.

The mixed methods research has pragmatism as theoretical underpinning [128]. As such, it allows a combination of methods from two different epistemological views on knowledge, the realist and the constructivist perspectives [128]. Pragmatism enables the researchers to draw on both quantitative and qualitative assumptions, and to choose the methods and

procedures that best meet with their needs and purposes [133]. However, it should not be interpreted as just a practical approach [128].

Many proposals for classification (typologies) of mixed-methods research have been published [130, 134]. Two important aspects for design and conduct of a study with mixed methods are the purpose of the study and the timing. Greene et.al. [129] classified five purposes: triangulation (seeks convergence with several investigations of the same phenomenon to strengthen the validity of the results); complementarity (seeks richer understanding by combining information from complementary sources or methods); development (results from one method informs the other method); initiation (generate new insights) and expansion (includes several outcomes e.g. processes and outcomes).

The sequence is another important factor [130, 132]. Data collection and analysis could be parallel or sequential. Guest proposed a focus on the points of interface between two datasets, i.e. where, how and why datasets are connected and mixed [130]. Points of interface could be at the sampling, data collection, data analysis and interpretation stages.

There are also several options for how to integrate qualitative and quantitative results. If data should be integrated during the analysis, data can be transformed. Qualitative data can be transformed into quantitative (“quantitizing”) or vice versa (“qualitizing”). Alternatively, qualitative and quantitative data can be presented together [135].

#### **GRADE framework**

The Grading for Assessment, Development and Evaluation, GRADE, framework [136] is an on-going development by an international working group of methodologists. GRADE is a classification system for the grading of evidence, originally developed for intervention studies. GRADE classifies evidence as high, moderate, low or very low.

GRADE is a tool to systematize the assessment of uncertainties in the body of evidence, usually a meta-analysis. It comprises five domains: methodological weaknesses (risk for bias), inconsistent results from the studies included in the meta-analysis (inconsistency), large confidence interval for the result of the meta-analysis (imprecision), problems with transferability of the result to the context of the research topic (indirectness) and risk for publication bias. The strength of the evidence is lowered if there are weaknesses in at least one of the domains.

#### **Framework for assessment of trustworthiness**

The rigour of findings in qualitative research can be judged with the help of sets of criteria. Several sets exist, underpinned by different scholar traditions [137]. Lincoln and Guba proposed a set of criteria, labelled trustworthiness [138]. Trustworthiness builds on social science and comprises four domains. Credibility (parallel to internal validity) refers to how well the methods used address the research question. Credibility includes choice of participants, selection of methods for data collection and richness of data. Member checks and negative case analysis increase the credibility. Transferability is a parallel to external

validity. Judgement of transferability requires that the study gives a sufficient description of e.g. context, and characteristics of the participants. Dependability (reliability) judges whether there have been any changes over time that may influence the findings. Confirmability (objectivity) finally, could be achieved by an external expert, who validates the results against the original data.

### **Use of theoretical underpinnings in the thesis**

Two behaviour change theories, the social cognitive theory [114] and the theory of planned behaviour [113], and the habit theory [115] were used to explain behavioural findings from paper 1.

Paper 2 was based on the GRADE framework for the assessment of evidence. There was an introduction on how GRADE can be used to assess the evidence for diagnostic accuracy [139], but there was little guidance how to apply the principles. Nor were any systematic reviews using GRADE published. Part of the work therefore was to develop practices how to apply GRADE.

Paper 3 used mixed methods and its pragmatic approach to design the study [129], collect and analyse data.

The intervention in paper 4 was based on the habit theory [115], and the analysis was based on the COM-B model [104].

## 2 RATIONALE AND AIMS

This thesis is the answer to two questions that evoke my curiosity as a project manager at SBU. The first question related to the diagnosis of depression, and emanated from the publication of the SBU report on treatment for depression in 2004 [140]. The report concluded that a diagnosis was essential for the adequate management of depression, and the use of instruments was advocated to support diagnosis. There were claims from the profession that such instruments deteriorated the communication with the patient and thus had a detrimental effect on the consultation. However, no studies were published that could confirm or deny those arguments. A knowledge gap thus was identified.

The second question related to the well documented gap between evidence and practice [141]. SBU reports were blamed for being mere shelf warmers. As responsible for activities to support uptake of the results of the reports, I became interested in whether the uptake could be enhanced.

My opportunity came in 2011, when I worked with a systematic review on instruments for screening, diagnosis and follow up of mood disorders [70]. It gave me a chance to merge my two fields of interest in one project: to work with the implementation of a SBU report on instruments for depression. However, as the results from the systematic review evolved, the scope for the implementation activity was broadened to depression and anxiety.

### Aims

The overall aim of the research was to create knowledge about which instruments for depression that are evidence based, and a strategy to support the use of such instruments in primary care.

### Specific study aims

Study 1 (Paper 1): to explore determinants of practice for the use of instruments for depression as part of the consultation in Swedish primary care

Study 2 (Paper 2): to determine the diagnostic accuracy for depression of case-finding instruments, structured and semi structured interviews and severity measures

Study 3 (papers 3 and 4): pilot study of the use of the MINI for patients at risk for depression or anxiety in Swedish primary care

- Paper 3: to explore whether the MINI is acceptable and useful as experienced by patients, FPs and therapists
- Paper 4: to explore whether an implementation strategy, where patients are referred to a therapist for the MINI, is feasible for patients, FPs and therapists, and to explore factors that influence the referral.



### 3 METHODS

This chapter describes the methods used, as well as ethical considerations. The thesis has an overarching before- after design and comprises four papers. The two first papers provided the answer to which instrument to choose, and a set of barriers and enablers in primary care to instruments in general. In a planning phase, not described in the papers, the implementation object (the instrument chosen) and the implementation strategy (intervention) were decided. The last two papers explored the feasibility of the instrument in a pilot trial. Table 4 presents an overview of the papers.

**Table 4** Overview of the papers in the thesis

Paper	Design	Objective	Data source	Data analysis
1	Qualitative	Explore determinants of practice for instruments	Focus groups with FPs	Systematic text condensation [142]
2	Systematic review	Find the implementation object	Systematic literature search	Meta-analysis
3	Observational	Explore acceptability of the implementation object	Semi structured interviews and focus groups Structured questionnaires	Inductive content analysis Descriptive statistics Triangulation
4	Observational	Explore acceptability of the intervention	Semi structured interviews and focus groups	Inductive and deductive content analysis

#### Design

Paper 1 is a qualitative interview study using focus groups to gather data.

Paper 2 is a systematic review. It was partly conducted at SBU, as their first diagnostic accuracy systematic review built on the GRADE. Paper 2 is an update which used another method for meta- analysis.

The mixed methods design for the study which is presented in papers 3 and 4 had a complimentary purpose with multiple informant groups: the patients, the physicians and the therapists.

## Setting

The qualitative studies (papers 1, 3 and 4) took place in primary care. It is the entry level for mental disorders in Sweden, except when the psychiatric emergency room is warranted. An FP examines the patient and agrees with the patient on the diagnosis and on further steps to be taken, e.g. a treatment or referral to secondary care. Patients have the right of being informed and to make decisions about their health and treatment together with the physician [143].

Patients sign up for a specific FP at a primary care centre (PCC) of their own choice [144]. However, for acute problems, they most probably will meet with another doctor, who does not know their personal situation and medical history. As in other countries, Swedish FPs work with tight schedules. Typically, a consultation for an acute problem lasts for 10 – 15 minutes.

Many PCCs have psychosocial teams with counsellors that support in e.g. crisis situations. As evidence suggests that psychotherapies give similar rates of improvement and recovery as anti-depressive drugs [140], this has fueled a demand for professionals certified in psychotherapies, mainly psychologists. However, there is a lack of psychologists and referral to external consultants has been common. It should be noted that even patients, who know or suspect that they have a mental problem and would like to have psychological treatment, must meet with an FP first.

## Settings for the studies

The studies were conducted in two health care regions, Västra Götaland (paper 1) and Stockholm County (papers 3 and 4). Stockholm County had separate guidelines for depression [145] and anxiety disorders [146]. In the depression guidelines, several instruments were compiled in a list without ranking [145]. The guidelines for anxiety disorders on the other hand, recommended a specific structured interview and specific screening instruments [146]. The PCCs also had access to [www.viss.nu](http://www.viss.nu), (VISS), a comprehensive website with links and recommendations. It was developed by the county to support primary care and was based on regional and national guidelines. VISS included links to a broad range of instruments for anxiety and depression, without notice of their evidence for accuracy. Västra Götaland had local guidelines for depression including use of some screening instruments, but no efforts had been made from the county to implement them.

In paper 1, the participants were recruited from two areas: the city of Gothenburg and its suburbs, and the middle-sized town of Skövde with rural surroundings. In Gothenburg, the FPs were recruited from an on-going RCT that evaluated whether the use of a self-rating questionnaire for severity of depression had any impact on patient outcomes [147]. All FPs

had undergone a ½ day training with an experienced colleague, but not all had been in the intervention group and used the questionnaire. The participants from Skövde were recruited from six PCCs by a colleague at the regional primary care research centre and had no previous training or education on instruments for depression.

In papers 3 and 4, seven PCCs were approached. However, two requirements for the study hampered participation. The PCC should have employed or contracted therapists, and furthermore there was no funding for the participation. Therefore, costs for the therapist time had to be borne by the PCC budget. Finally, two PCCs agreed to participate. One FP moved to another PCC during the study and continued to recruit patients at the new workplace, PCC3. PCC1 had no previous experience of structured interviews. PCC2 (and thus the FP at PCC3) had already implemented the MINI, and the FPs were trained to use the interview. The PCCs are described in Table 5.

**Table 5** Characteristics of PCCs included in papers 3 and 4

<b>PCC ID</b>	<b>Location</b>	<b>Listed number of patients</b>	<b>Psychosocial burden (CNI*)</b>	<b>Number employed physicians**</b>
1	Suburb	18 000	CNI = 1.26	15
2	Suburb	21 000	CNI = 0.93	14
3***	Central Stockholm	10 000	CNI = 0.72	3

\* CNI = Care Need Index [148] a measure of psychosocial burden, where higher values indicate larger problems; average CNI = 1.0; \*\* under training or family physicians; \*\*\* not included in paper 4

### The implementation object

The object for implementation in papers 3 and 4 was the SBU report [70], or rather, a piece of evidence. SBU found that a structured interview, the MINI International Neuropsychiatric Interview (MINI) [149], had 95 percent sensitivity and 84 percent specificity for depression. Recalling that the sensitivity of an FP is around 50 percent, use of the MINI could represent an opportunity to improve the detection rate of depression.

The MINI is a comprehensive instrument. At the time of the studies, the MINI was based on DSM-IV (MINI 6.0) and captured 15 psychiatric diagnoses. The MINI is constructed in sections, one section per diagnosis. They comprise questions that can only be answered with “yes” or “no”. Many questions are supplemented with examples to facilitate the understanding of the question. Each section starts with some questions about core symptoms. If patients answer “no” to them, the remaining questions in the section are skipped. Questions deal with time frame, duration and severity of symptoms included in the criteria. Each section

has a final check box, where the interviewer notes whether criteria are fulfilled or not. After having completed the sections, the interviewer judges which is the primary diagnosis (if any).

The MINI thus gives information about a range of psychiatric disorders. SBU had only evaluated the MINI for depression and bipolar disorder. However, studies have also measured the diagnostic accuracy for other disorders, and found acceptable accuracy for panic disorder and generalized anxiety disorder in psychiatric and primary care settings [149-151]. The accuracy for agoraphobia and social anxiety disorders has been evaluated in psychiatric settings and is acceptable [149, 151]. Therefore, we broadened the intended use of the MINI to support detection of depression and anxiety. This better mirrors the population in primary care, where differentiation of depression and anxiety or between anxiety disorders is a common issue.

### Implementation strategy

The implementation strategy, described in papers 3 and 4, was based on the literature, including paper 1, and the previously described survey about use of instruments, conducted by SBU [70] (See Background). The literature on family physicians' perceptions about instruments is limited to case-finding and severity measures for depression [38, 78-81, 85]. Box 3 describes the determinants for practice regarding instruments for mental problems identified from these sources, as well as from two open-ended questions in the SBU questionnaire (results not published).

The conclusion was that implementation of the MINI needed a strategy that focused on existing habits, lack of time, and knowledge and skills about use of the MINI.

#### *Barriers*

- The MINI takes long time
- Prefer to rely on own clinical experience
- Prefer to work according to own, established routines
- Printed questionnaires do not fit with the FP professional role and consultation style
- The consultation is disturbed by introducing questionnaires and important information is lost
- The patient has problems with the questionnaires

#### *Enablers*

- Facilitates communication with complicated patients
- Advocated by trusted colleagues
- Evidence that patients and the health care benefit from the use

**Box 3** FPs' determinants of practice in primary care for use of instruments for mental problems

Habit theory suggests that interventions that focus on changes in the context that maintains the habit have a greater probability of success [115]. A task shift could be one way to circumvent the habits, and also meet the barrier lack of time. As the ultimate goal was that FPs used the results of the MINI as part of their diagnostic process, the actual interview could be performed by someone else. The MINI is developed to be used by any medical professional after short training [149]. We therefore chose to investigate a task shift where the FPs could refer patients for a MINI assessment and have the results fed back afterwards. Therapists were chosen as they are educated and trained to use tests and instruments. However, the FPs could choose whether they wanted to conduct the MINI by themselves or refer to a therapist.

Supportive education of the FPs was the other part of the strategy. The FPs were invited to two introductory sessions at their clinic. One was about the value of correct diagnosis and the role of the MINI. The second session dealt with the evidence for the MINI.

Two of the therapists had not used the MINI. They got one full day education and hands-on training with an experienced FP. The third therapist already knew the MINI and had no training.

## Participants

Family practitioners (papers 1, 3 and 4), patients (papers 3 and 4) and therapists (papers 3 and 4) participated in the data collection.

## Staff

Table 6 shows characteristics of the FPs and therapists that participated.

**Table 6** Characteristics of family practitioners and therapists that contributed in focus groups or interviews.

	<b>Paper 1, FPs</b>	<b>Papers 3 and 4, FPs</b>	<b>Papers 3 and 4, therapists</b>
Total number (women)	27 (17)	17 (14)	3 (2)
≥ 20 years as certified FP/therapist	10	7	1
0-19 years as certified FP/therapist	15	7	1
Under specialist training	2	3	1

## Patients

Papers 3 and 4 targeted patients who had symptoms that could suggest depression or anxiety. We applied criteria from a study on screening for depression [152]. Patients could seek medical advice for a new episode of mental health problems, or have somatic symptoms that could not be explained in biological terms. Frequent attenders were also

eligible, as frequent consultations have been associated with symptoms of depression and anxiety [152-155]. We added another group to the criteria: patients with an on-going depression that had not responded adequately to treatment after two months. Patients had to be 18 years old or older, sufficiently capable in the Swedish language to understand and answer the questions in the MINI, and without cognitive deficiencies. Patients, who were in acute need of treatment, were excluded.

## Data collection

The methods used for data collection comprised questionnaires (papers 3 and 4), semi structured interviews (papers 3 and 4), focus groups (papers 1, 3 and 4), and a systematic literature search followed by a systematic selection of studies (paper 2).

## Questionnaires

Paper 3 used a set of published questionnaires, one to the patients and one to the therapists. The questionnaires had been developed and validated in a German study [156] on the acceptance of a structured interview, Diagnostisches Interview bei Psychischen Störungen (DIPS for the DSM-IV-TR). The questionnaires were translated into Swedish language and back-translated to German language by a native German living in Sweden. The accuracy of the back-translated text was verified by one of the head researchers in the German study.

The questionnaires had a key question, which measured the global satisfaction with being interviewed with the MINI or to interview with the MINI. The degree of satisfaction was estimated with a Visual Analogue Scale (VAS, 0 = not at all satisfied to 100 = totally satisfied). Ten Likert-type statements, scored from 0 = do not agree at all, to 3 = agree fully, captured factors relevant to the therapist-patient relationship in two areas: mental effort and emotional reactions. Items dealt with if the interview was exhausting, intrusive or too comprehensive, if a good relationship was established and the therapist understood the patients and its problems, and if the interview was of help for the patient. The therapist was also asked to rate if he or she felt competent enough during the interview, items that fell outside the two areas.

For paper 4, a question was added to the patient questionnaire. The patients were asked to rate the global satisfaction with being referred to a therapist.

The questionnaires also gathered information about the patients' reasons for visiting the PCC, the duration of the assessment, gender and age of the patients, and the therapists' interpretation of the MINI results.

The questionnaires were completed directly after the MINI assessment, put in a coded envelope and sealed by the informant. In total, 125 patients consented and completed the questionnaire while another seven did not consent. One patient did not return the questionnaire. Some patient characteristics are listed in Table 7.

**Table 7** Characteristics of patients included in papers 3 and 4.

<b>Characteristics</b>	<b>PCC1</b>	<b>PCC2</b>	<b>PCC3***</b>	<b>Total</b>
<b>Number of patients (% female)</b>	55 (72 %)	54 (83 %)	16 (81 %)	125 (78 %)
<b>Age distribution*</b>				
< 25 years	5	4	4	13 (12 %)
25 – 60 years	41	31	12	84 (78 %)
> 60 years	9	5	0	14 (13 %)
<b>MINI diagnosis, current disorder**</b>				
Depression only	10	5	7	22 (18 %)
Depression and anxiety disorder (s)	11	3	7	21 (18%)
Anxiety disorders only	14	23	0	32 (27 %)
Other MINI diagnoses, with depression and/or anxiety	6	3	2	11 (9%)
Other MINI diagnoses	1	2	0	3 (2 %)
None	13	13	0	26 (22%)

Notes: \* age missing for 14 patients; \*\* MINI diagnoses missing for five patients, \*\*\* not included in paper 4

### Semi structured interviews

Semi structured face-to face interviews were conducted with patients and therapists to get a deeper understanding of experiences and perceptions. For patients, a purposeful sample was sought. The goal was to obtain an even distribution between PCCs, and a variation with respect to gender, age, ethnicity and occupation.

All patients who completed the questionnaire were asked to participate, and 49 volunteered to be contacted for more information. Finally, 24 persons consented and participated in an interview. Nine did not consent, four did not turn up for the interview and 16 women (25 – 60 years) from PCC2 were not contacted in order to maintain a balance. The sample comprised six men and 18 women, 13 from PCC1, eight from PCC2 and three from PCC3. Of these, four were younger than 25 years, and five were older than 60 years. Six interviewees were first- or second-generation immigrants.

The interviews were conducted concurrently, which in practice meant two weeks to two months later than the questionnaires and took an average of 25 minutes. They were audio recorded.

The interview aimed at capturing experiences and perceptions of the steps of the clinical pathway and of the MINI. A topic guide contained broad questions of type “what happened when you met with your doctor the first time?”. There were additional probes to be used if necessary, e.g. “how do you perceive that your confidence in your doctor is after the proposal

to refer for the MINI?”. After the first three interviews, the guide was modified. Two questions about which profession was best suited to make a psychiatric diagnosis, and conduct the MINI were added.

The two therapists at PCC2 were interviewed together and the therapist at PCC1 was interviewed individually, at their respective offices. The interviews were conducted at study end, lasted for 45 minutes and were audio recorded. The guide covered topics on their experiences of the MINI and how they believed that the patients perceived the interview, as well as cooperation and team work with the FPs.

### **Focus groups**

Focus group discussions were chosen for FPs in papers 1, 3 and 4. Focus groups is a technique that was developed in consumer research [157-159] but is increasingly being used in health care research [157]. A moderator presents topics and the participants are encouraged to discuss freely with each other. The interaction between the participants can reveal perspectives that the researchers have not considered. The role of the moderator, apart from introducing new themes for discussion, is to ensure that all group members get the opportunity to talk. An observer helps by writing notes and supporting the moderator. Focus groups should have an appropriate size for discussions, between four and eight are recommended [157], a recommendation that was followed.

The discussions followed topic guides with key questions, and probing questions to be given if they were not addressed spontaneously. For paper 1, the key questions were inspired by the previous studies on use of instruments [78, 79], and dealt with experiences of using various types of instruments, how the consultation is affected by their use and reliability of the instruments. After the first two group discussions, a fourth key question regarding previous educations on use of instruments was added as this had been brought up as important.

Topics in paper 3 and 4 dealt with the process, from the first patient meeting to the decision of treatment, and a broad area about advantages and disadvantages with the MINI. The focus groups were conducted when recruitment of patients had stopped.

### **Systematic literature search and selection of studies**

For paper 2, literature was retrieved from searches in four electronic data bases up to April 29, 2014.

The following inclusion criteria were applied:

- Clinical populations (not epidemiological studies)
- Conducted in Europe, North America, Australia or New Zealand
- Evaluated the diagnostic accuracy at the established threshold
- Semi structured or structured interviews as reference test
- Maximal time between the index and reference test one week

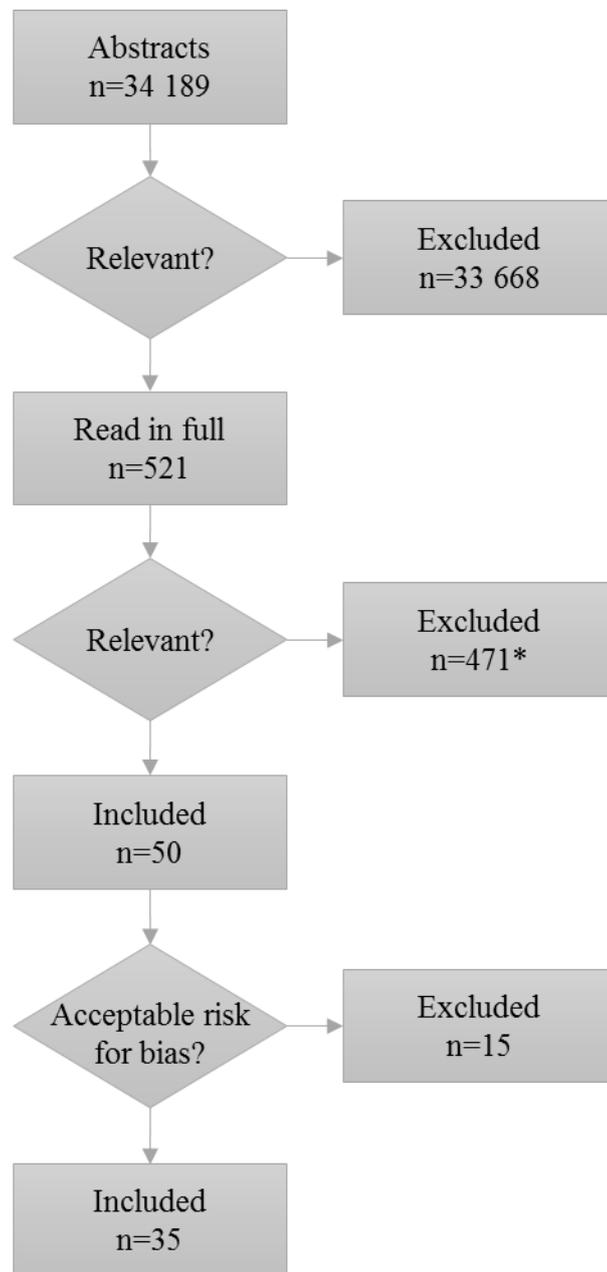
- Full-text articles published in peer-reviewed journals in the English, Scandinavian, German and French languages

The systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [160].

Historically, diagnostic accuracy studies have been characterized as of weak design and superficial reporting of conduct and results [161]. The Standards for Reporting of Diagnostic Accuracy, STARD, statement pinpoint the importance of a relevant reference standard, a relevant study population and acceptable time between tests [161]. The risk of bias was assessed using the Quality Assessment of Diagnostic Accuracy Studies, QUADAS, checklist [162] which mirror the items of STARD. QUADAS comprises 11 items. It captures problems with selection of the population and the choice of reference standard as well as assessment of design and procedures (e.g. blinding, drop-out, handling of missing data). Only studies with a low or moderate risk of bias were included in the meta-analyses.

Thresholds for acceptable minimum sensitivity and specificity were defined: 80 percent sensitivity and 80 percent specificity for structured and semi structured interviews and 80 percent sensitivity and 70 percent specificity for case-finding instruments. This is lower than proposed by e.g. Gilbody et.al. [163] but is based on the understanding that instruments are only a part of the diagnostic procedure.

The flow chart for inclusion of studies is illustrated in Figure 3. Ultimately 35 studies were used for the analyses.



**Figure 3** Flow chart for selection of studies in the systematic review, according to PRISMA (160). \* Reasons for exclusion: modified index test (n = 54); reference standard not an interview (n = 105); time between tests (n = 41); data for contingency tables not reported (n=31); not an accuracy study (n=230)

## Data analyses

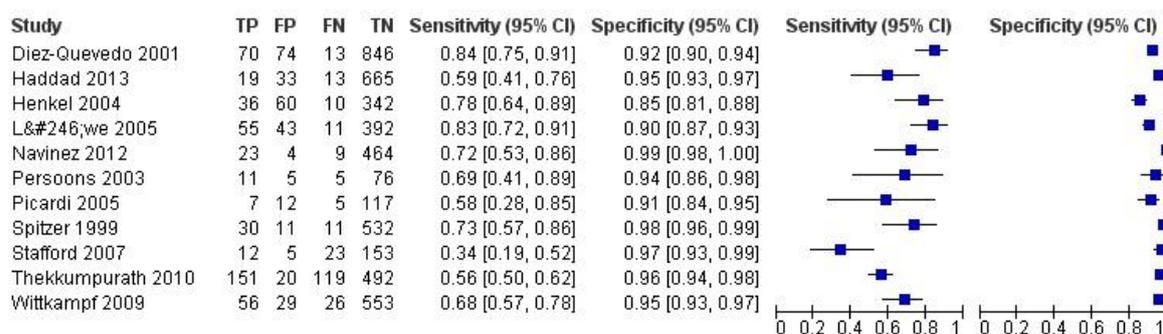
### Quantitative analyses

In papers 3 and 4, the VAS scores, for the global satisfaction of use of the MINI and for referral to a therapist, were estimated with a ruler. The scores had a right-skewed distribution and therefore the medians and interquartile range (IQR) were calculated. The proportion of participants who fully or almost fully agreed with the statement (rated 2 or 3) was calculated for the ten items on the acceptance questionnaires.

The time required for the MINI was analysed as the mean (SD) number of minutes.

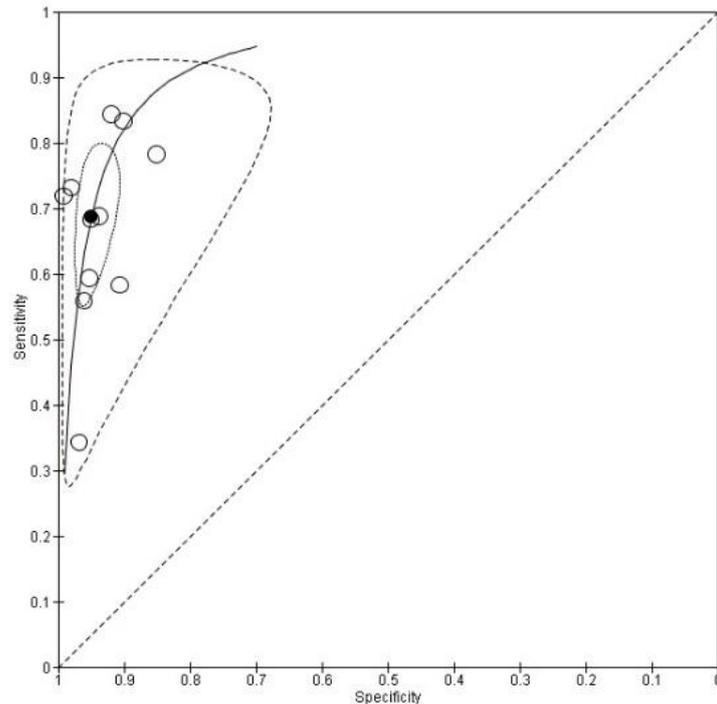
### Hierarchical meta-analysis for diagnostic accuracy

Earlier, the average sensitivity and specificity was calculated with univariate methods. They resulted in two separate meta-analyses, one for sensitivity, and one for specificity (“paired forest plots”), see Figure 4 for an example. Thus, the correlation between sensitivity and specificity was not taken into account, which could lead to an underestimation of the accuracy [164].



**Figure 4** Example of a paired forest plot meta-analysis (the sensitivity and specificity for algorithm-based PHQ-9). No average is calculated in this example

Nowadays, two other methods are recommended instead of the paired forest plots [165, 166], the bivariate model [167] or the Hierarchical Summary Receiver Operating Curve, HSROC model, [168]. Both models operate at two levels, within studies and between studies. HSROC and bivariate models have different underlying assumptions. The bivariate model assumes that there is one threshold (e.g. colour of a dipstick) and results in an average sensitivity and specificity for this threshold. The HSROC, on the other hand, assumes that there are several thresholds. The analysis results in an average across thresholds. However, if no covariates are entered into the meta-analysis, the methods yield the same results [169]. Figure 5 shows an example of a bivariate meta-analysis.



**Figure 5** Example of a bivariate meta-analysis (algorithm-based PHQ-9). The sensitivity and specificity for each study is represented by a white dot, and the average sensitivity and specificity by a black dot. The inner area depicts the 95 % confidence region and the outer area the 95 % prediction area. The curve shows the correlation between sensitivity and specificity.

Paired forest plots could easily be created in e.g. the Cochrane collaboration software RevMan. The bivariate and HSROC models require statistical expertise, and experience from programs such as STATA or SAS [165]. A practical solution, used in paper 2, is to generate the values in a statistical program and enter the data into RevMan for a graphical representation (see Figure 5). For paper 2, an external statistician applied the METADAS macro for SAS [170]

Bivariate or HSROC analysis require that at least four studies are available [165]. For several of the instruments in the systematic review, only two studies were included. In these cases, the Cochrane Handbook [165] recommended paired forest plots. These analyses were run with the MetaDiSc [171] software.

### Systematic text condensation

Paper 1 used systematic text condensation [142]. This is a straightforward method, developed to be used by beginners without expertise in qualitative methodology.

Systematic text condensation is inspired by Giorgi's psychological phenomenology [172], and has a similar four-step procedure. The first step was to get an overview of data. Next, the relevant units of meaning were identified. They were assembled in code groups, that were split into subgroups. The meaning units within each subgroup were resynthesized to an

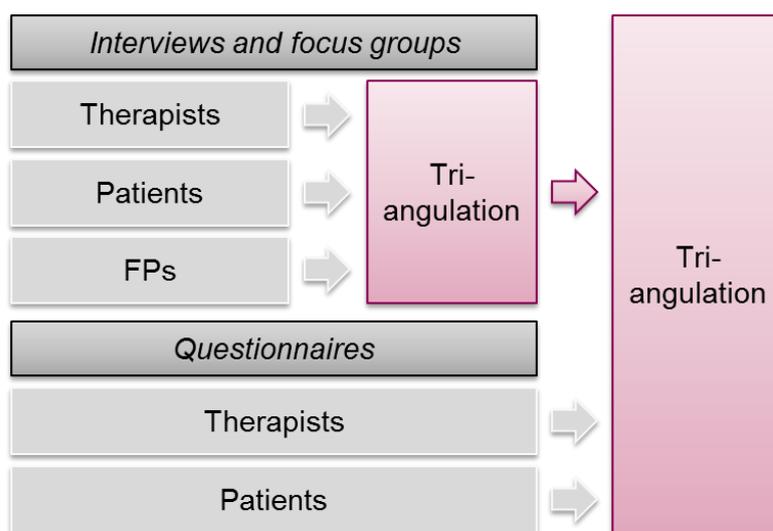
artificial quotation. Finally, the contents of the quotations were synthesised back to the code group level as a story, and concentrated into a category heading.

### Content analysis

For papers 3 and 4, qualitative content analysis was used [173, 174]. Paper 3 was based on an inductive analysis [173, 175], i.e. no codes were defined a priori. In paper 4, inductive analysis was followed by deductive, where the COM-B components were used as main categories. For both papers, the categories were triangulated across participant groups. The process is described in Figure 6.

### Triangulation as mixed method

We integrated the qualitative and quantitative results in paper 3 by triangulation. Triangulation, originally a navigation term, refers to a simple method to determine the position using observations from two additional points. Triangulation can be used to improve validity of findings and to explore complementarity [176]. A mixed methods matrix was created where items from the questionnaires were matched to one of the main categories from the content analysis [135]. Thus, neither quantitative nor qualitative findings were transformed.



**Figure 6** Analysis, papers 3 and 4

### Ethical considerations

Research should be conducted with respect for the participant individuals, and their integrity should be protected. The two studies (papers 1, paper 3-4) were granted ethical approval and were carried out in accordance with the Helsinki declaration [177].

The participants were supplied with an information folder about the study, its data handling and publication intents. They signed informed consent after reading. They were informed about their right to withdraw at any time without any consequences to their treatment and that their contribution would be deleted. Results were only reported at group level and quotations were coded in order to ensure confidentiality. For papers 3 and 4, the patients were asked to

participate in the research after they had decided to accept the MINI-interview. Thus, there were some patients that went through the MINI but declined to participate in the data collection.

There were actual and potential ethical threats to the participating patients, i.e. in the study described in papers 3 and 4. The threats were related to the MINI and to the research procedure. The MINI as well as the procedures may represent a threat to equality. The version of the MINI used in the study is in the Swedish language, and the study included only patients that had sufficient understanding of the language. This criterion had practical reasons since the MINI, although translated and validated for several languages, are not available for languages common with immigrants to Sweden. A possibility would be to use interpreters, but then there is a risk that the validity of the instrument is threatened by layman translation. So, if the MINI-assessment has a patient benefit, patients with difficulties speaking or understanding Swedish, would be denied this advantage.

There was a risk that the MINI could cause harm. Patients might feel offended or upset by the MINI-interview as it probes for symptoms that can be difficult to face and talk about. However, the study that dealt with a similar interview [156] did not indicate that negative emotions should be a problem. Furthermore, the MINI was conducted by therapists or FPs with experience of handling emotional situations.

The research instruments, the questionnaires and the interview, on the other hand were less problematic, as they asked for experiences of using the MINI without references to the patients' problems.

The study could represent a threat to autonomy. However, patients which might have a lowered capacity to make an informed consent were excluded, such as patients with cognitive deficiencies or deep depression. Finally, an interview represents a situation with unequal power between the interviewer and the interviewee. In order to minimize the threat, the participants could choose the time and venue for the interview, and efforts were made to make the interview comforting and pleasant.

For FPs and therapists, the ethical threats were considered to be very small.

## 4 RESULTS

### Determinants for use of depression instruments in primary care

Paper 1 found that most determinants for the use of depression instruments in primary care relate to the knowledge and motivation of the individual FP. However, other stakeholders could have an influence.

Six determinants were identified: the agenda is set outside primary care, instruments seldom add value for the FP, the dialogue with the patient suffers, the scores cannot be trusted, instruments do not fit primary care and patient-rated instruments are valuable in specific situations. The results can be categorized with COM-B, as shown in Table 8.

**Table 8** Determinants for use of depression instruments among Swedish primary care physicians

<b>Determinant</b>	<b>Component in COM-B</b>	<b>Acts as enabler</b>	<b>Acts as barrier</b>
Properties of the instruments	Mental capability (knowledge)	Evidence that instruments give benefits	Scores cannot be trusted
Self-efficacy	Mental capability (knowledge)		Own knowledge sufficient for detection and diagnosis Not trained in how to integrate instruments in the consultation
Influence from other agencies	Social opportunity	Trusted colleagues advocate their use	Industry marketing of instruments to support sales of antidepressants
Beliefs about relationship with the patient	Reflective motivation		Belief that the use of instruments decreases the patients' confidence in the competence of the FP

Continued next page

Beliefs about consequences to the FP	Reflective motivation	Facilitates communication with some patients (e.g. somatising or silent patients)	Hampers the dialogue Takes time from other, more important, tasks Gives a bureaucratic consultation style and makes the job more boring
Beliefs about patient benefits	Reflective motivation	Scores can facilitate sick leave compensation and intake to psychiatric care	Not suitable for vulnerable patients with psychiatric problems
Established habits	Automatic motivation		Prefer to work in the traditional way (“ingrained in the walls”)

---

### Theories can help to explain the findings

Classic theories were used to explain the behaviours. In summary, FPs preferred to rely on their knowledge and clinical experience, and continue to work without support of instruments. These findings are consistent with self-efficacy from the Social Cognitive Theory [114] and habit formation [115]. However, sometimes the FPs decided to use a case-finding or severity measure, if this would help the patient. Use of this tactic can be explained by the Theory of Planned Behaviour [113].

### The diagnostic accuracy for depression was sufficient for three instruments

The included studies in the paper 2 systematic review concerned structured and semi structured interviews and case-finding instruments. We did not retrieve any relevant studies regarding instruments to measure severity.

Many instruments were evaluated in one study only and were judged to have a very low strength of the evidence (insufficient, according to the SBU terminology). The average sensitivity and specificity and the strength of the evidence for instruments with at least two studies are summarized in Table 9. As seen from Table 9, only three of them, the Structured Clinical Interview for DSM-IV-Axis 1 (SCID) [178], the Mini International Neuropsychiatric Interview (MINI) and the Patient Health Questionnaire- 9 (PHQ-9) [179], had evidence supporting a sensitivity and specificity above the pre-set benchmarks.

SCID is a semi structured interview, which needs certified interviewers and typically lasts for 1,5 hours. As such it is less suitable for primary care. The PHQ-9 at cut-off 10 could be an alternative. However, in the original SBU report, there were fewer studies on PHQ-9 and

these were highly heterogeneous [70]. SBU concluded that PHQ-9 lacked evidence. Therefore, the target for implementation became the MINI.

**Table 9** Summary of findings for diagnostic accuracy of instruments for depression with structured interviews as reference [180]

<b>Instrument</b>	<b>Outcome</b>	<b>Average (95% CI)</b>	<b>Strength of evidence</b>	<b>Above threshold</b>
SCID-I	Sensitivity	86 (73–94)	Low	Yes
	Specificity	92 (88–95)	High	
MINI	Sensitivity	95 (93–97)	High	Yes
	Specificity	84 (80–87)	High	
PRIME-MD	Sensitivity	<70	Moderate	No
	Specificity	85 (82–88)	High	
BDI-II, cut-off 14	Sensitivity	92 (83-97)	Moderate	No
	Specificity	72 (58-82)	Very low	
HADS-D, cut-off 7	Sensitivity	70 (55-82)	Low	No
	Specificity	83 (73-90)	Low	
PHQ-9, algorithm	Sensitivity	69 (60-76)	Moderate	No
	Specificity	95 (92-97)	High	
PHQ-9, cut-off 10	Sensitivity	88 (77-94)	Moderate	Yes
	Specificity	78 (65-88)	Moderate	
CES-D, cut-off 16	Sensitivity	95 (83-99)	Very low	No
	Specificity	33-73 %	Very low	

### The MINI is useful and well accepted by patients and staff

Paper 3 described the experiences of using the MINI as perceived by the referring FPs, the interviewers (therapists and three FPs), and the patients.

Six main categories emerged from the analyses. Three categories concerned strengths and weaknesses of the MINI. Another two categories captured consequences of using the MINI, and the sixth category dealt with the role of the MINI in the diagnostic process.

## **Strengths and weaknesses of the MINI**

In general, characteristics of the MINI, in terms of overall structure and comprehensibility of questions, were appreciated by patients and interviewers. A perceived drawback was that some problems, that are common in primary care, were not addressed, e.g. stress (interviewers) and subthreshold diagnoses (patients). The format, whereby patients could only answer yes or no, was not appreciated (although some patients preferred it).

The questions were not offensive or intrusive to the patients and few were exhausted by the interview.

The length of the MINI assessment was acceptable to patients and therapists, but too lengthy for FPs that conducted the MINI.

## **Results from the MINI can give benefits to FPs and patients**

An advantage for the FPs was that the diagnoses became more accurate. Co-morbidities that could influence treatment effects were discovered, as well as disorders with sensitive or awkward symptoms. The MINI was perceived as a standard test for psychiatric complaints, analogous with routine tests for e.g. diabetes.

For the patients, the MINI assessment could give new insights about their problems. Furthermore, it was a relief to get a diagnosis. It could help patients cope with the situation, read and think about how to proceed. Ultimately, the results of the assessment could have an impact on the treatment.

## **The MINI is just one part of the diagnosis**

Patients were not always clear about the purpose of the MINI. Some of them seemed to believe that the MINI should replace the clinical assessment. Many patients underscored that a MINI interview should be combined with a conversation to deepen the understanding of the problems. The interviewers agreed that an interpersonal contact was essential before arriving at a diagnosis.

## **Referral to a therapist for MINI assessment can be feasible in primary care**

The second part of the study of the MINI in primary care was described in paper 4. At two PCCs, the FPs got a new opportunity to refer patients to an in-house therapist.

Nine out of 15 FPs at PCC1 referred at least one patient (average 7, range 1 to 17) for a MINI assessment. At PCC2, the behaviour was not anticipated. The FPs earlier had conducted the MINI themselves. During the study time, they dropped the MINI and made a basic clinical examination. Patients in need of more comprehensive examination were referred to a therapist for diagnosis and treatment. The FPs and the therapists at PCC 2 stated that referral for MINI assessment was not of interest to them.

Several factors were identified that could influence the FP referral to a therapist, as seen in Table 10.

A conclusion from the study is that the task shift could be feasible in primary care, as seen at PCC1. However, factors such as the competence of the FPs and how professional roles are defined between FPs and therapists will affect whether the process is appropriate or not, as seen at PCC2.

The patients appreciated to be referred to a therapist for the MINI assessment. They expressed that therapists in general had a higher competence to handle psychiatric problems as well as to conduct the MINI. However, knowledge, experience and personality was of more importance than the professional education, and an FP could be as well equipped to conduct MINI as a therapist.

**Table 10** Factors related to competence, opportunities and motivation that can influence the FP probability to refer

<b>Factors that <i>increase</i> the probability to implement referral</b>	<b>Factors that <i>decrease</i> the probability to implement referral:</b>
The FP perceives good knowledge and skills in psychiatry and judges that the patient needs a comprehensive assessment	The FP perceives good knowledge to assess the patients' problems without referral
The FP perceives insufficient knowledge and skills and appreciates support in the diagnostic work	
Perceptions that the FP consultation times are too short for patients with psychiatric problems	There is a lack of routines for e.g. feedback of results
Easy access to therapists	
Process facilitator or change agent available	
Perceptions that the information given by the patient in the MINI assessment is more reliable when conducted by a therapist, as the patient feels less hurried and the environment is more comforting	The FP perceives that the patient is unwilling to make an extra visit or to see a therapist
The FP perceives that the work load is eased	
A teamwork with the therapist is established which is experienced as rewarding and fun	
Early, positive experiences of benefits of the referral	Early, negative experiences of the referral Easier to work according to established routines

## 5 DISCUSSION

### Discussion of the results

Identification of barriers and facilitators to the use of instruments were essential parts of the thesis. Papers 1 and 3 had quite different contexts in terms of the FPs' education and knowledge. In paper 1, the FPs had limited experience from depression specific questionnaires and none from the MINI. Most of their education on instruments had been organized by the pharmaceutical industry. In paper 3, the FPs were educated by a recognized champion and colleague. Around half of these FPs already had own experience from conducting the MINI themselves. The other half got experience from referring the patients for an assessment.

Accordingly, many determinants were different. The FPs in paper 1 perceived that instruments most often disturbed the consultation. The FPs in paper 3 appreciated the comprehensive information gathered from the MINI. To them, the MINI contributed to a more accurate diagnosis and hence to a more adequate treatment. Thus, with proper training and support, the MINI can be useful in primary care.

However, some important barriers were similar between the papers. The FPs expressed that often their own knowledge was sufficient to make a diagnosis – without the MINI. Furthermore, even if they knew that the MINI could be advantageous for the patients, it was easy to fall back on their established routines. Thus, not even the opportunity for referral, was always sufficient to overcome those routines.

Beliefs about the patient were another common barrier. Patients were considered fragile and should not be exposed to standardized questionnaires. However, the patients appreciated the standardization. They also perceived that it was a relief to answer questions instead of telling their stories in own words. Thus, patients may be overrated as barriers.

### Determinants for use of instruments for depression in primary care

Paper 1 identified six determinants, which can influence the FPs' use of instruments for detection and diagnosis of depression. One determinant related to external actors, while the remaining concerned knowledge and motivation of the individual physician. Most of the barriers have been described in studies, primarily from the UK [57, 78-81]. The enabler, that instruments were good communication tools for some patients, has been described in a study from the US [181]. Scores from instruments were useful for communication with actors such as the Swedish Social Insurance Agency (Försäkringskassan) and psychiatric care. However, the strength of this enabler depends on local requirements and agreements and may be of less importance outside Västra Götaland.

## Choice of instrument

Paper 2 found that only three instruments fulfilled our benchmark criteria, two structured interviews, SCID-I and the MINI, and a patient-rated case finding instrument, PHQ-9 at cut-off 10.

Few studies have investigated the diagnostic accuracy of structured interviews. Only two studies on SCID were retrieved, and the evidence that SCID has sufficient sensitivity for depression is low. As SCID is an established reference standard, the lack of studies is surprising. The MINI on the other hand, is relatively well investigated. Apart from the two included studies, three others were identified. Two were excluded due to the geographical setting [150, 151], and one due to high risk for bias [182]. However, these studies gave similar results and the evidence base for the MINI seems to be stable.

There is reason to feel more concern about PHQ-9. The results from paper 2 agreed with those reported in two other systematic reviews with broader inclusion criteria [183, 184]. However, PHQ-9 is still a subject for new studies, which rapidly make the systematic reviews obsolete. Two recent systematic reviews [185, 186] had a high heterogeneity in the meta-analyses, and several studies were outside the 95 percent prediction region. One of these systematic reviews included a subgroup analysis. This showed that the sensitivity varied between settings, being 81 percent for primary care and only 70 percent for secondary care [186]. Thus, the last word regarding PHQ-9 has not been said, and the evidence for its diagnostic accuracy is not clear cut.

### The MINI was acceptable and useful in primary care

Paper 3 showed that the MINI was appreciated by FPs, patients and therapists.

The use of structured interviews has been controversial, both in psychiatric and primary care, with proponents as well as opponents. In paper 3, most physicians were positive and found the results of the MINI useful in their management of the patients. This confirms results from other studies on the feasibility of the MINI, from primary care in Brazil [150] and psychiatric care in Norway and Italy [187, 188].

The physicians perceived that the MINI helped to establish the diagnosis needing treatment as well as (unknown) comorbidities that could influence the choice of treatment. Some physicians also noted that they might be too prone to label patients as depressed without further investigations. The MINI thus could be a help to avoid misdiagnosis. These findings agree with the literature [189-193].

The time required to conduct MINI was seen as a problem by the physicians that had experience from using it. Patients on a non-scheduled 10 minutes' consultation had to book another visit, which often was not realistic. The lack of time is cited as a reason for not using structured interviews in primary care [150, 194].

In the study, most patients were satisfied with the MINI. They saw many advantages and very few reported negative experiences. No other studies that dealt with primary care patients' experiences of structured interviews were identified. However, the findings correspond with results from studies in other settings, mostly psychiatry [156, 187, 188, 195-197].

The format, where the respondent is limited to answer "yes" or "no", is probably the major source of complaint, from patients as well as interviewers. Interviewers with more experience of the MINI added open-ended questions when they needed additional information. Thus, in practice the MINI was sometimes used as a semi-structured interview. This is acceptable according to the instructions for the MINI [149] but requires greater skills and more in-depth training of the interviewer.

Depending on context it can be feasible to refer patients to a therapist for MINI assessment

Paper 4 showed that a task shift, where therapists conducted the MINI and fed back the results to the physician, was feasible at one PCC but not in the other.

We anticipated that FPs would be reluctant to use the MINI in their consultation. However, the actual target behaviour was that the FPs should use the results from the MINI. Therefore, a referral for the assessment was explored. For somatic disorders, referral to other specialists for more comprehensive examinations and tests are routine. Referral for psychiatric structured interviews is not routine. It has been tried in some studies, where educated nurses assessed patients with SCID-I or the Global Mental Health Assessment Tool, GM-HAT, [198]. The referral process was well accepted by physicians [69, 199]. No studies where patients were referred to other professionals for a structured interview were retrieved.

FPs at both PCCs used the opportunity to refer, although the purpose at PCC2 was diagnosis and treatment. A common denominator for the PCCs was that the physicians' competence in psychiatry influenced the referral. Many FPs perceived a good knowledge in diagnosis of mental disorders, in line with current literature [38, 39]. They reserved referral for complicated patients, while FPs with lower self-perceived competence referred at a broader scale.

A difference between the two centres was the presence of a process facilitator or change agent. At PCC1, the therapist voluntarily took the role as facilitator. At PCC2, the former manager had a goal about a better management of patients with psychiatric problems, which included an assessment with the MINI for complicated patients. As the manager left, no one else continued the change process, and the physicians returned to their old routines. The value of a facilitator or change agent has been shown in several studies, e.g. based on the PARIHS framework [200]. Presumably, the absence of such a function contributed to the return to consultations as usual.

## Methodological considerations

The approach for the papers was pragmatic, where the research methods were chosen to answer the research questions. The designs chosen were a systematic review (paper 2), a mixed methods study (paper 3) and two interview studies (papers 1 and 4). It has to be said that a study with qualitative methods only, is not the best design for paper 4. A randomized study where MINI with referral was compared to MINI without referral, and with interviews as part of the data collection, had given a more comprehensive picture. However, such a study was not possible to conduct.

### Paper 2

The systematic review was conducted according the principles of PRISMA [160], which was a strength of the paper. The included studies were assessed for risk of bias with the validated QUADAS check list. Although two researchers independently rated the studies before agreeing on a final decision there is always a matter of subjectivity in the ratings. As studies with high risk of bias were not included in the meta analyses, the rating played a larger role than in many other systematic reviews.

Results of a systematic review always depend on the predefined selection criteria as studies not fulfilling criteria are filtered out. One criterion, that had a heavy impact, was that studies should be conducted in countries with similar cultures and beliefs about depression [201]. In hindsight, it may be questioned if this exclusion criterion was relevant. On the one hand, a study conducted in the Netherlands and Surinam supported that cultural differences could create bias [202]. On the other hand, large systematic reviews do not indicate that results from other parts of the world systematically deviate from results in e.g. Europe and the US [185, 186]. In practice therefore, other systematic reviews should have better power in the meta analyses.

An important issue is if the result of a meta-analysis is relevant. Studies may be so heterogeneous that an average of their results hardly is meaningful. Heterogeneity can result from e.g. differences in the study population and choice of reference standard. For questionnaires, additional issues refer to the translation to other languages, and that the meaning may be understood differently. An example is the HADS item “butterflies in the stomach”, which was hard for Arabic people to understand and relate to [203]. Given the large heterogeneity seen in the meta-analyses for PHQ-9, it can be questioned whether there really is a meaningful average across settings and patients. The way forward may be, as indicated by Moriarty et.al. [186], to conduct fine grained meta-analyses, based on assumptions on how depression is perceived by different patient groups.

### Papers 1, 3 and 4 – qualitative methods

Systematic text condensation and qualitative content analysis was used. This was a reasonable choice, since the aim was descriptive and not explanatory.

A strength was that conduct and reporting of the studies were guided by the QOREC statement for qualitative studies [204]. Other strengths and limitations are discussed from the Lincoln & Guba criteria for trustworthiness [138].

## **Paper 1**

### *Credibility*

The study recruited physicians from personal networks of two senior FPs. The number of participants in the focus groups was acceptable [157, 158] and they had a mix in terms of age, gender and experience of working as family practitioners. A possible limitation was that the focus groups took place after work hours and so, the participants may have been unbalanced in terms of interest in the research question. Yet, the groups covered a spectrum of opinions; negative to instruments, ignorant about them or positive.

A risk with focus groups are that participants do not share their real experiences and perceptions. Two reasons are that they may want to please the researchers or want to avoid conflicts with other participants [158]. However, the participants did neither seem to hesitate to express conflicting opinions, nor to attempt to please the researchers.

### *Transferability*

All participants had stable employments and most knew their patients well. They worked at primary care centres that did not hire temporary doctors. These working conditions may have impacted their routines and ways of relating to the instruments, limiting transferability to other primary care centres. As the focus groups discussed from the perspective of voluntary use of instruments, some of the findings may not be relevant to countries where use of instrument is encouraged or mandatory.

### *Dependability*

The focus groups were conducted within a period of four months, and no events occurred that might have impacted the views of the participants. The topic guide was modified though, to include effects of education. This change would presumably not have affected the findings.

### *Confirmability*

The transcripts, codes and analyses were read by a senior researcher and family practitioner in order to validate the findings.

## **Papers 3 and 4**

### *Credibility*

This was a convenience sample of primary care centres, which is a limitation. However, the participating FPs had a variation in terms of age, gender and years as family practitioner.

The patients were recruited consecutively and the sample was varied in terms of age, gender, occupation and ethnicity.

A problem might be that the amount of data is not sufficient to capture variations [205, 206]. Therefore, a reevaluation of sample size was conducted, as suggested by Malterud et.al. [206]. As a result, more interviews were conducted until the last two interviews did not yield any new information.

The data collection from the FPs was not complete. Everyone was invited to the focus groups, but almost all of those that participated had used the MINI or referred patients. Thus, we have no information from the physicians that were not interested in the MINI. There may be more barriers against the use of the MINI or a referral than those that we captured.

A member check was conducted with therapists and physicians at one PCC, but the other PCCs and the patients were not offered opportunities to comment on the analyses.

### *Transferability*

The sampling method may have led to limitations in the transferability of the results. The PCCs were interested enough to invest in the study, as there was no financial support. The experiences and perceptions of interviewers and GPs may differ in PCCs with other values and priorities, although patient perceptions likely are more stable.

Furthermore, the experiences related to the DSM-IV version of the MINI, which has been replaced by the DSM-5 version. This would probably not influence the findings. However, it has been discussed to introduce a web version of the MINI. A change of format may lead to other perceptions from both the patients and the interviewers.

### *Dependability*

The study was conducted during one full year, and there might be a risk that the experiences changed as a result of internal or external events. The FPs at PCC2 changed their perceptions about use of the MINI from the introductory education meeting to the focus groups discussions, as a result of time constraints and new routines. No shift over time was observed for the other FPs or the therapists.

On the other hand, it became apparent that some patients had changed their perceptions from the time when the questionnaire was completed, to the research interview, which was conducted several weeks later. At the interview, the patients had most often been informed of their diagnosis, had started treatment, and had also had time for reflection which may have affected their views. However, this probably did not alter the findings, as there was no systematic change. Patients expressed higher as well as lower acceptance at the interview.

### *Confirmability*

This was a limitation since the findings were not validated by an external expert.

### Paper 3, questionnaires

We used a published, validated questionnaire to measure satisfaction with the MINI and the interview situation. The data collection had limitations. As the study coordinator at one of the PCCs left, the study got a lower priority. This affected the inclusion of patients and the completion of therapist questionnaires negatively. However, there was little variability in the completed 115 questionnaires. Thus, there is no reason to believe that the missing questionnaires would have substantially changed the results.

### Considerations on choice of determinants and theory

The study in paper 4 used three of the four steps considered necessary to design an intervention [207]. Barriers (or determinants of practice) were identified, these were linked to behaviour change interventions, and the feasibility was tested on the target groups. However, theory did not inform the data collection.

### **Appropriate determinants of practice?**

The determinants were gathered from the international literature, from a Swedish focus study and a national survey. The desired behaviour was *use of the results from MINI*, rather than use of the MINI. The determinants were appropriate for PCC1 but not wholly appropriate for PCC2. Time became an overriding determinant for the FPs at PCC2, while knowledge about psychiatry and skills in conduct of the MINI, became of less importance.

### **Appropriate behaviour change interventions?**

With use of the terminology of BCW [100], the components training, modelling and environmental restructuring (task shifting) were chosen. The central component was the task shifting. This was appropriate for PCC1 but not for PCC2. The referral for an assessment only was not interesting for the physicians, who trusted their own assessments more. Furthermore, the therapists objected as they saw the MINI as an integrated part of their treatment. The evidence for task shifting is very limited and no published studies have explored shifts between FPs and therapists [117].

### **Appropriate theoretical underpinning for the analysis?**

Determinants that related to individual factors were selected. Thus, determinants related to interactions with actors outside the PCCs were excluded. Therefore, frameworks as the CFIR or the framework for implementation in primary care [121] were of limited use.

An interesting alternative, which was not considered, is the Normalization Process Theory, NPT [127]. The NPT comprises four domains: coherence, cognitive participation, collective action and reflexive monitoring, and has an organizational perspective. In retrospect, the NPT might have taken the perceptions of the therapists into account in a better way. Factors, such as co-operation between the therapists and the physicians, and the importance of routines and communication policies would then have had more weight in the conclusions. The NPT has

also been advocated as one of a limited number of frameworks that are suitable for chronic care settings [208].

## 6 CONCLUSIONS

This thesis showed that

- there were several determinants of practice that influenced when, and to what extent Swedish FPs used instruments for depression. These mainly concerned factors related to the knowledge and attitudes of the individual physician. However, some actors outside primary care could influence the use.
- the MINI is one of few instruments with good diagnostic accuracy for depression.
- the MINI was appreciated by physicians, patients and therapists in primary care. The time for the MINI assessment could be a problem for physicians who want to conduct the MINI.
- A task shift, where the physicians refer patients to a therapist for the MINI and the results are fed back to the physician, can be feasible in primary care. Whether a task shift with a therapist is useful depends on contextual factors.

## 7 CLINICAL IMPLICATIONS

The thesis focused on FPs and patients. This is a simplified model as many parts of the inner and outer context were not investigated. The role of e.g. national guidelines and agreements for care thus remains unclear. Yet, the findings have the potential to have clinical implications.

One example is the new National guidelines for depression and anxiety, issued by the National board for health and welfare 2016 [209]. The guidelines underscore the importance of a correct diagnosis and recommend use of the MINI. The backdrop for the guidelines is that the use of instruments probably has not changed substantially since the SBU survey 2011. This is suggested from a web based survey by the Board for health and welfare in late 2016 [210]. Few respondents used structured interviews while instruments to assess severity were more common. However, the survey was launched at a website restricted to members of the Swedish College of General Practice (SFAM), and only 76 physicians responded. As most counties were represented in the survey, a reasonable conclusion, is that the SBU report has had very little impact and that there still may be reluctance against the MINI. Paper 3 therefore could be used as a support for the implementation of the guidelines.

A practical issue is who should conduct the assessment. Paper 4 showed that referral, in this case to a medical social worker, certified in CBT, worked well. Other studies have indicated that referral to a trained nurse may be another option. A third option is that the FPs conduct the MINI themselves. A fourth option is that the FPs refer patients with suspected mental problems to a psychologist for diagnosis and treatment. This option was used by one of the centres in paper 4. For alternatives 1-2, implementation of the MINI is a strategic decision as it is associated with new tasks for the staff and new routines.

## 8 FUTURE RESEARCH

The original aim of the thesis was to develop an implementation strategy for an evidence based method from a SBU report, and to evaluate it in a randomized trial. However, the plan was not realized. One reason was that the evidence based method, the MINI, lacked feasibility data for primary care. Therefore, the aim was modified to study feasibility in a pilot trial, with a prospect of these findings being used for future trials.

The thesis contributes with a tiny fragment of knowledge regarding the overarching issue of management of mental problems in primary care. The papers build on an assumption that a correct diagnosis affects the prognosis and wellbeing of the patients. However, even if there is some data supporting this assumption, it still has to be proven. Randomized trials are needed that allocate patients to diagnosis with the MINI or to routine clinical assessment, and effects on treatment choice and depression outcomes are measured.

The version of the MINI in the thesis was based on paper and pen. There were complaints from some interviewers that the “bunch of paper” was impractical and heavy to use. They had preferred it as an interactive web-based version to simplify their work. Such a version is provided by the developers. However, it is only available in English language and it is not free of charge. It could be worthwhile to negotiate with the developers to get a national license for the web- based version, translate it and evaluate the feasibility.

The strategy chosen lacked a component that we had identified from paper 1: that the motivation to use an instrument would increase if the higher management (e.g. managers at county level) recommended it, and also followed up the use and fed back comments. It would have been interesting to evaluate this component, e.g. by studying perspectives at meso and macro levels. This would require a far larger study, ideally from several counties [211].

## 9 SVENSK SAMMANFATTNING

Denna avhandling handlar om två vanliga problem i primärvården, depression och ångest; diagnostik som en viktig del i omhändertagandet; samt en strategi som stöd för att införa evidensbaserad kunskap från en SBU-rapport till svensk primärvård. SBU-rapporten handlade om diagnos av depression. Avhandlingen avgränsades ursprungligen till att gälla depression men breddades till att även omfatta ångestsyndrom.

De ingående studierna kom fram till att:

- det fanns många hinder för att använda bedömningsformulär som stöd vid handläggning av patienter med misstänkt depression i primärvården. Läkarna uppfattade sig ha tillräcklig kompetens i psykiatrisk diagnostik, och de hade utformat egna rutiner baserat på klinisk erfarenhet. Formulären störde samtalet med patienten, passade inte in i rutinerna och stämde inte med deras syn på läkarrollen.
- en strukturerad intervju, Mini International Neuropsychiatric Interview (MINI), har tillräcklig diagnostisk tillförlitlighet för depression för att vara kliniskt relevant som del i en utredning av psykisk ohälsa i primärvården.
- MINI uppskattades av patienter och läkare för att bidra till en korrekt diagnos och en adekvat behandling. Tidsåtgången kunde vara ett problem för läkare som genomförde MINI.
- remittering till en psykolog eller kurator för MINI, där resultatet återkopplas till läkaren, kunde vara ett sätt att underlätta användningen av MINI. Läkarnas motivation för att remittera berodde på flera faktorer, bland annat egen kompetens i psykiatrisk diagnostik, arbetsbelastning och om det finns en förändringsagent i organisationen.

### Introduktion

Cirka 1/3 av patienterna i primärvården har någon form av psykisk ohälsa, där nedstämdhet och ångest är de vanligaste problemen. Det är viktigt att kunna skilja ut dem som har en diagnos så att de får tillgång till den behandling som fungerar bäst. Det finns ett visst stöd för att rätt diagnos har en positiv påverkan på patientens besvär och tillfrisknande.

Diagnostiken blir därmed en väsentlig del av omhändertagandet.

Idag finns inga biokemiska markörer för ångest och depression. Diagnosen avgörs utifrån i vilken utsträckning patienten uppfyller kriterierna i något av de diagnostiska klassificeringssystemen. Grunden för diagnostiken är samtalet mellan läkare och patient, kompletterat med undersökningar för att utreda om problemen kan ha någon kroppslig orsak.

Systematiska översikter har visat att knappt hälften av patienter med depression eller ångest i primärvården identifieras. Samtidigt har andra studier pekat på en överdiagnostik av framförallt depression. Det finns ett visst stöd för att diagnostiken kan förbättras om läkarna använder bedömningsformulär som ett komplement i konsultationen.

Bedömningsformulären är antingen upplagda som strukturerade intervjuer, där en kliniker ställer frågorna, eller frågeformulär som patienterna fyller i (skattningsskalor). Intervjuerna täcker ett större antal psykiatriska diagnoser medan frågeformulären avser en diagnos. Det finns många bedömningsformulär men deras tillförlitlighet har varit oklar.

Bedömningsformulär används i begränsad omfattning i svensk primärvård. I enkäter har ca en tredjedel av läkarna svarat att de använder patientskattade formulär regelbundet, medan ett litet fåtal använder strukturerade intervjuer.

Studier har visat att det är svårt att införa nya, evidensbaserade metoder i primärvården. Stödaktiviteter (implementeringsstrategier), som är baserade på analyser av vilka barriärer och underlättande faktorer (determinanter) som finns, antas underlätta införandet.

## Syften

Avhandlingen syftade till att besvara följande frågor:

### *Fas 1*

Hur ser läkare i svensk primärvård på att använda bedömningsformulär för depression? Vilka faktorer påverkar användningen? (artikel 1)

Vilka bedömningsformulär har en tillräcklig diagnostisk tillförlitlighet för depression för att vara kliniskt relevant att använda? (artikel 2)

Baserat på artikel 2 valdes bedömningsformuläret MINI ut. Artikel 1 användes för att utforma en strategi som stöd för att införa MINI. I strategin hade läkarna möjlighet att remittera patienter till en psykolog eller kurator för MINI-bedömning.

### *Fas 2*

Är MINI acceptabelt och användbart i svensk primärvård för att upptäcka depression och ångestsyndrom? (artikel 3)

Är remittering till psykolog eller kurator för MINI-bedömning acceptabel och användbar? Vilka faktorer påverkar läkarnas benägenhet att använda MINI i kombination med stödstrategin? (artikel 4).

## Beskrivning av ingående artiklar och deras resultat

### Artikel 1

Studien använde kvalitativ metodik. Den genomfördes i Västra Götaland. Deltagarna rekryterades antingen från en pågående studie i Göteborg eller via förfrågan från primärvårdens forskningsenhet i Skövde. Sammanlagt genomfördes fem fokusgrupper med totalt 27 läkare. De hade jämn fördelning i antalet yrkesverksamma år, från ST-läkare till över 20 års erfarenhet, och nära 2/3 var kvinnor. Diskussionerna spelades in, skrevs ut och analyserades med hjälp av metoden systematisk textkondensering

Studien bekräftade delvis resultat från studier genomförda i andra länder. Läkarna diskuterade huvudsakligen skattningsskalor, som de flesta hade någon erfarenhet av att använda. Sammantaget litade läkarna mer på sin egen kunskap och kände att de själva inte hade någon nytta av resultaten från bedömningsformulär. Formulären passade inte in i deras rutiner och störde konsultationen. Speciellt strukturerade intervjuer kändes främmande att lägga in i samtalet och var dessutom tidskrävande.

Till skillnad från läkare i tidigare publicerade studier såg läkarna fördelar med att använda skattningsskalor i specifika situationer. De kunde fungera som hjälp att kommunicera med vissa patienter, och höga poäng på en skattningsskala kunde underlätta intag på psykiatrisk avdelning eller sjukskrivning.

Läkarna ansåg att om de fick kunskap om att bedömningsformulär fyllde någon nytta och träning i att använda formulären skulle det underlätta användning.

## Artikel 2

Detta var en systematisk litteraturöversikt som täckte såväl strukturerade intervjuer som formulär som var specifika för depression. Översikten var en uppdatering av ett kapitel i SBU-rapporten. Den genomfördes i enlighet med internationella riktlinjer för systematiska översikter (PRISMA).

Några viktiga avgränsningar var att studierna skulle vara genomförda inom primär- eller specialistvård i Europa, Nordamerika, Australien eller Nya Zeeland, att formulärens egenskaper jämfördes med strukturerade intervjuer som referens och att det fick gå högst en vecka mellan mätning med formuläret respektive referensen. Studiernas kvalitet bedömdes med en validerad checklista. Studier med god eller acceptabel kvalitet ingick i analyserna. Metaanalyserna gjordes med bivariat metodik och resulterade i medelvärden för känslighet och träffsäkerhet.

För att ett formulär skulle vara kliniskt relevant som tillägg till en rutinundersökning skulle det ha minst 80 procent känslighet och minst 70 procent (skattningsskalor) eller 80 procent (strukturerade intervjuer) träffsäkerhet. Endast tre formulär uppfyllde kravet: de strukturerade intervjuerna Structured Clinical Interview for DSM-IV Axis I disorders (SCID-I) och MINI, samt självskattningsformuläret Patient Health Questionnaire med nio påståenden (PHQ-9).

## Mellanspel

Avsikten med avhandlingen var att studera införande av en slutsats från den ursprungliga SBU-rapporten i primärvården. I rapporten hade endast MINI och SCID-I vetenskapligt stöd. MINI valdes som implementeringsobjekt eftersom SCID-I tar längre tid och kräver högre kompetens på bedömaren. MINI är accepterat och sett som användbart av såväl patienter som läkare i studier från psykiatrin. Däremot saknades motsvarande studier från primärvården.

MINI skulle bli mer användbart om utvärderingen gällde både depression och ångestsyndrom. Ångestsyndrom ingick inte i SBU:s rapport men studier som utvärderade MINI tog oftast upp tillförlitligheten för samtliga ingående tillstånd. I dessa studier hade MINI acceptabel tillförlitlighet för ett eller flera ångestsyndrom.

Resultaten från artikel 1 tydde på att läkarnas egen kunskap och invanda rutiner spelade störst roll för om de skulle använda bedömningsformulär. Utbildning skulle förmodligen inte vara tillräckligt för att börja använda MINI. Enligt en teori om vanebildning har organisatoriska förändringar som stör vanan störst chans att leda till ett ändrat beteende. Därför utformades en strategi där läkarna kunde välja att remittera patienter till en kurator eller psykolog (beteendevetare) och få återkoppling om resultatet av MINI. I den tredje studien undersöktes därmed såväl erfarenheter och uppfattningar om MINI som om att remittera patienter till en kurator eller psykolog för MINI-bedömning. Studiens resultat delades upp i två artiklar (3 och 4).

### Artikel 3

Studien ägde rum på tre vårdcentraler i Stockholmsområdet. Patienterna som ingick kunde ha psykiska problem, besöka vårdcentralen mycket ofta, ha kroppsliga symtom som inte kunde förklaras eller ha en diagnosticerad depression som inte hade förbättrats av behandling.

Datainsamlingen bestod av enkäter från 125 patienter och från dem som genomförde MINI (beteendevetare eller läkare,  $n = 115$  enkätsvar), semistrukturerade intervjuer med 24 patienter och tre beteendevetare, tre fokusgrupper med totalt 17 läkare samt mätning av tid för bedömningen. Enkäterna var publicerade och validerade för en annan, liknande, strukturerad intervju

Intervjuer och fokusgrupper analyserades med kvalitativ innehållsanalys, först separat för patienter, bedömare och läkare. Därefter analyserades resultaten över deltagargrupperna (trianglering) och slutligen matchades svaren från enkäterna till fynden från intervjuerna.

Analysen visade att såväl patienter som läkare och beteendevetare uppskattade MINI. MINI uppfattades bidra till en bättre behandling för patienterna. Förutom huvuddiagnoser var en fördel att samsjuklighet med andra psykiatriska diagnoser upptäcktes. Patienterna kände att de standardiserade frågorna hjälpte dem att få nya insikter och att våga erkänna även pinsamma symtom som tvång och riskbruk av alkohol. Få patienter kände att MINI var för omfattande eller utmattande eller att frågorna var kränkande. Å andra sidan var själva formatet en nackdel och många patienter hade svårt att bara svara ja eller nej på frågorna. MINI tog i medeltal 26 minuter att genomföra, vilket sågs som ett problem av läkare som hade erfarenhet av att genomföra MINI.

### Artikel 4

Läkaren på en av de tre vårdcentralerna genomförde MINI själv, och patienter från den vårdcentralen ( $n = 3$ ) uteslöts från analysen. Läkarna på de båda andra vårdcentralerna fick

nya möjligheter att remittera patienter till psykolog eller kurator. Resultaten baserades på intervjudata enbart och dessa analyserades med kvalitativ innehållsanalys.

En första analys undersökte hur den diagnostiska processen gick till. Den visade att den första vårdcentralen hade infört MINI på det sätt som var tänkt. Sextio procent av läkarna hade använt sig av remitteringen och såväl patienter som läkare och beteendevetare var tillfredsställda. Läkarna på den andra vårdcentralen hade däremot valt en icke förutsedd väg. Efter en rutinundersökning skickades patienter som behövde mera utredning vidare till psykolog för fortsatt handläggning inklusive bedömning med MINI och behandling.

En modell för beteendeförändring, Competence Opportunity Motivation -Behaviour (COM-B), användes för att identifiera faktorer som påverkade läkarnas benägenhet att remittera patienter till en beteendevetare. Analysen gjordes för samtliga patienter, oavsett om remitteringen gällde MINI-bedömning enbart eller diagnostik följt av behandling.

Bristande tid var en nyckelfaktor som talade för remittering. Konsultationstiden kunde vara för kort för att få tillräcklig information. Dessutom blev remittering till psykolog en avlastning i arbetsbörda för läkarna. Läkarens självupplevda kompetens i psykiatrisk diagnostik spelade roll. Läkare med hög tilltro till sin kompetens remitterade utvalda patienter med komplexa problem. Läkare som var mera osäkra remitterade även mindre komplicerade patienter. Korta väntetider för besöket hos beteendevetare och rutiner som stödde remitteringen var andra viktiga faktorer. Många läkare upplevde dock att det var lätt att falla tillbaka i gamla vanor.

## Kommentarer

Studierna har visat att MINI har förutsättningar att användas i svensk primärvård. De bidrar därmed med en skärva av kunskap om hur diagnostik av psykiska sjukdomar kan förbättras. Fortfarande kvarstår en central fråga: huruvida användning av MINI leder till snabbare förbättring eller tillfrisknande. Till detta behövs randomiserade studier.

## 10 ACKNOWLEDGEMENTS

Things take time. This thesis is the result of a long learning journey, and a dream since nearly 45 years has finally come true. I am deeply grateful to many, who have supported me with Competence, Opportunities and Motivation. I want to thank my employer, **SBU**, for the generosity to give me time for my studies. **Måns Rosén** approved and encouraged my idea to follow a SBU-report, and supported me as supervisor. My managers, **Susanna Axelsson** and **Charlotte Hall**, allocated time and gave me the feeling that my research was a priority. I would also like to extend my gratitude to **Nina Rehnquist**, who once employed me and trusted that a person from marketing in the pharmaceutical industry could be transformed into a governmental agency systematic reviewer.

Supervisors! It has been comforting to lean against you and your experience and to have access to your wise guidance. Many thanks to **Ingvar Krakau** for all your efforts, especially as main supervisor during the last year. You were always at my disposition, “just a click away”. I will remember with joy the garden consultations for problem-solving and new ideas. Your extensive network of contacts in primary care and in the Stockholm county has been invaluable to my research. And, at times, your infinite sources of philosophical and fun quotations from the literature, or links from YouTube were restorative. To **Rolf Wahlström**, who contributed with scientific rigour and critically reviewed plans and manuscripts – and also shared my affection for Austin roses. To **Eva-Lisa Petersson** who guided me through my first stumbling steps in qualitative methodology. The focus group discussions were a great learning experience and we had fun when travelling around to primary care centres. To **Sonja Modin**, for your great patience in teaching me the art of content analysis and keeping me on track when I made too large jumps to the conclusions. I have appreciated those Face Time meetings where we could solve various issues.

Co-authors and more! A warm thank you to all of you. To **Kristina Boström Bengtsson** and **Cecilia Björkelund** who also were essential for study 1 as you recruited participants for the focus groups study. **Sandra af Winklerfelt Hammarberg**, we have shared ups and downs and we have had great discussions during the years. I wish you all luck in your own PhD studies. Maybe we will find the time to write that article about the readiness-to-change questionnaire that became a leftover from my thesis. **Henna Hasson**, you have been a source of inspiration long before our paths crossed. Your presentation at a seminar, about how you used a theoretical framework, was an eye-opener. Later on, you have generously taken your time to guide me in the jungle of theories, frameworks and models. **Petter Gustavsson**, we have known each other since we worked on the SBU- report and you have kindly been available for advice and knowledge about validation of instruments. **Berit Dybäck**, you have been instrumental as coordinator for the MINI study with your dedication, energy and feeling for order. I have really appreciated working together with you.

Colleagues and friends have been there when I needed them. **Jenny Odeberg**, **Karin Stenström**, **Annika Sahlin Platt** and **Christina Edlund**: your continuous encouragement

and hints to the expectations for the party afterwards have been a push at times. **Agneta Brolund** volunteered as a trouble shooter for the EndNote program and conducted the update literature search for the systematic review. A warm thank you for your support. To the **PROCOME team**: thank you for opening your door to an external student, sometimes lost in the practical routines of KI. Being part of a group and listening to others' experiences and knowledge have meant much to me.

And for the final touch: thank you to **Jeff Platt** for editing the English and to my son **Calle** for the layout work and technical support.

This thesis would not have been written, had it not been for some very special persons: my family and my mentor, **Lisa Ekselius**. Lisa, you have always had a listening ear. You have, as a psychiatrist, given me insights in how to think about diagnosis but more importantly you have helped me to sort out practical and administrative problems. I have really appreciated to have you at my side. **Lennart**, you have shown a never-ending patience for all those weekends and holidays spent with a computer rather than with you. I am also deeply thankful for the more substantial support in form of bread baked, laundry fixed, house cleaned and cats cared for. And to all of you, Lennart, Calle and **Lasse**, thanks for believing that I could make it!

## 11 REFERENCES

1. Joling KJ, van Marwijk HW, Piek E, van der Horst HE, Penninx BW, Verhaak P, et al. Do GPs' medical records demonstrate a good recognition of depression? A new perspective on case extraction. *J Affect Disorders*. 2011;133:522-7.
2. Abed Faghri NM, Boisvert CM, Faghri S. Understanding the expanding role of primary care physicians (PCPs) to primary psychiatric care physicians (PPCPs): enhancing the assessment and treatment of psychiatric conditions. *Ment Health Fam Med*. 2010;7:17-25.
3. World Health Organization (WHO)/Wonca. Integrating mental health into primary care: a global perspective. Geneva. 2008.
4. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness: the chronic care model, Part 2. *JAMA*. 2002;288:1909-14.
5. Bishop TF, Ramsay PP, Casalino LP, Bao Y, Pincus HA, Shortell SM. Care Management Processes Used Less Often For Depression Than For Other Chronic Conditions In US Primary Care Practices. *Health Aff (Millwood)*. 2016;35:394-400.
6. American Psychiatric Association (APA). Diagnostic and Statistical Manual of Mental Disorders (DSM-5). Washington: American Psychiatry Press; 2013.
7. World Health Organization (WHO). The International Statistical Classification of Diseases and Related Health Problems, ICD-10. 5th edition. World Health Organization; 2015.
8. American Psychiatric Association (APA). Diagnostic and statistical manual of mental disorders. 4th edition. Washington: American Psychiatry Press; 1995.
9. Sadek N, Bona J. Subsyndromal symptomatic depression: a new concept. *Depress Anxiety*. 2000;12:30-9.
10. Pincus HA, Davis WW, McQueen LE. 'Subthreshold' mental disorders. A review and synthesis of studies on minor depression and other 'brand names'. *Br J Psychiatry* 1999;174:288-96.
11. Penninx BW, Nolen WA, Lamers F, Zitman FG, Smit JH, Spinhoven P, et al. Two-year course of depressive and anxiety disorders: results from the Netherlands Study of Depression and Anxiety (NESDA). *J Affect Disorders*. 2011;133:76-85.
12. Conradi HJ, Bos EH, Kamphuis JH, de Jonge P. The ten-year course of depression in primary care and long-term effects of psychoeducation, psychiatric consultation and cognitive behavioral therapy. *J Affect Disorders*. 2017;217:174-82.
13. Steinert C, Hofmann M, Kruse J, Leichsenring F. The prospective long-term course of adult depression in general practice and the community. A systematic literature review. *J Affect Disorders*. 2014;152-154:65-75.
14. Demyttenaere K, Bruffaerts R, Posada-Villa J, Gasquet I, Kovess V, Lepine JP, et al. Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. *JAMA*. 2004;291:2581-90.
15. Lotfi L, Flyckt L, Krakau I, Martensson B, Nilsson GH. Undetected depression in primary healthcare: occurrence, severity and co-morbidity in a two-stage procedure of opportunistic screening. *Nordic J Psychiatry*. 2010;64:421-7.
16. Anseau M, Dierick M, Buntinx F, Cnockaert P, De Smedt J, Van Den Haute M, et al. High prevalence of mental disorders in primary care. *J Affect Disorders*. 2004;78:49-55.

17. King M, Nazareth I, Levy G, Walker C, Morris R, Weich S, et al. Prevalence of common mental disorders in general practice attendees across Europe. *Br J Psychiatry* 2008;192:362-7.
18. Mergl R, Seidscheck I, Allgaier AK, Moller HJ, Hegerl U, Henkel V. Depressive, anxiety, and somatoform disorders in primary care: prevalence and recognition. *Depress Anxiety*. 2007;24:185-95.
19. Bandelow B, Michaelis S. Epidemiology of anxiety disorders in the 21st century. *Dialogues Clin Neurosci*. 2015;17:327-35.
20. Lejtzén N, Sundquist J, Sundquist K, Li X. Depression and anxiety in Swedish primary health care: prevalence, incidence, and risk factors. *Eur Arch Psychiatry Clin Neurosci*. 2014;264:235-45.
21. Roca M, Gili M, Garcia-Garcia M, Salva J, Vives M, Garcia Campayo J, et al. Prevalence and comorbidity of common mental disorders in primary care. *J Affect Disorders*. 2009;119:52-8.
22. Martin-Merino E, Ruigomez A, Wallander MA, Johansson S, Garcia-Rodriguez LA. Prevalence, incidence, morbidity and treatment patterns in a cohort of patients diagnosed with anxiety in UK primary care. *Family practice*. 2010;27:9-16.
23. Badescu SV, Tataru C, Kobylinska L, Georgescu EL, Zahiu DM, Zagrean AM, et al. The association between Diabetes mellitus and Depression. *J Med Life*. 2016;9:120-5.
24. Gili M, Comas A, Garcia-Garcia M, Monzon S, Antoni SB, Roca M. Comorbidity between common mental disorders and chronic somatic diseases in primary care patients. *Gen Hospital Psychiatry*. 2010;32:240-5.
25. Khawaja IS, Westermeyer JJ, Gajwani P, Feinstein RE. Depression and coronary artery disease: the association, mechanisms, and therapeutic implications. *Psychiatry (Edgmont)*. 2009;6:38-51.
26. de Groot M, Anderson R, Freedland KE, Clouse RE, Lustman PJ. Association of depression and diabetes complications: a meta-analysis. *Psychosom Med*. 2001;63:619-30.
27. Chang WH, Lee IH, Chen WT, Chen PS, Yang YK, Chen KC. Coexisting geriatric anxiety and depressive disorders may increase the risk of ischemic heart disease mortality-a nationwide longitudinal cohort study. *Int J Geriatr Psychiatry*. 2016; Dec 14. doi: 10.1002/gps.4646
28. Chesney E, Goodwin GM, Fazel S. Risks of all-cause and suicide mortality in mental disorders: a meta-review. *World Psychiatry*. 2014;13(2):153-60.
29. Lumme S, Pirkola S, Manderbacka K, Keskimaki I. Excess Mortality in Patients with Severe Mental Disorders in 1996-2010 in Finland. *PLoS One*. 2016;11:e0152223.
30. Osby U, Brandt L, Correia N, Ekblom A, Sparen P. Excess mortality in bipolar and unipolar disorder in Sweden. *Arch Gen Psychiatry*. 2001;58:844-50.
31. Britton PC, Bohnert KM, Ilgen MA, Kane C, Stephens B, Pigeon WR. Suicide mortality among male veterans discharged from Veterans Health Administration acute psychiatric units from 2005 to 2010. *Soc Psychiatry Psychiatr Epidemiol*. 2017 Apr 11. doi: 10.1007/s00127-017-1377-x
32. Läkemedelsverket. Läkemedelsbehandling av depression, ångestsyndrom och tvångssyndrom hos barn och vuxna. Information från Läkemedelsverket; Uppsala; 2016.

33. National Institute for Health and Clinical Excellence (NICE). The treatment and management of depression in adults (update). Clinical guideline 90. National Institute for Health and Clinical Excellence; 2009.
34. National Institute for Health and Clinical Excellence (NICE). Generalised anxiety disorder and panic disorder in adults: management. CG 113. <https://www.nice.org.uk/guidance/cg113>. National Institute for Health and Clinical Excellence; 2011.
35. National Institute for Health and Clinical Excellence (NICE). Obsessive-compulsive disorder and body dysmorphic disorder: treatment, CG 31. <https://www.nice.org.uk/guidance/cg312005>. National Institute for Health and Clinical Excellence; 2005
36. National Institute for Health and Clinical Excellence (NICE). Post-traumatic stress disorder: management, CG 26 <https://www.nice.org.uk/guidance/cg262005>. National Institute for Health and Clinical Excellence; 2005
37. National Institute for Health and Clinical Excellence (NICE). Social anxiety disorder: recognition, assessment and treatment, CG 159. <https://www.nice.org.uk/guidance/cg159>. National Institute for Health and Clinical Excellence; 2013.
38. Barley EA, Murray J, Walters P, Tylee A. Managing depression in primary care: A meta-synthesis of qualitative and quantitative research from the UK to identify barriers and facilitators. *BMC family practice*. 2011;12:47.
39. Schumann I, Schneider A, Kantert C, Lowe B, Linde K. Physicians' attitudes, diagnostic process and barriers regarding depression diagnosis in primary care: a systematic review of qualitative studies. *Family practice*. 2012;29:255-63.
40. Andersson SJ, Troein M, Lindberg G. Conceptions of depressive disorder and its treatment among 17 Swedish GPs. A qualitative interview study. *Family practice*. 2001;18:64-70.
41. Piek E, Nolen WA, van der Meer K, Joling KJ, Kollen BJ, Penninx BW, et al. Determinants of (non-)recognition of depression by general practitioners: results of the Netherlands Study of Depression and Anxiety. *J Affect Disorders*. 2012;138:397-404.
42. van Rijswijk E, van Hout H, van de Lisdonk E, Zitman F, van Weel C. Barriers in recognising, diagnosing and managing depressive and anxiety disorders as experienced by Family Physicians; a focus group study. *BMC family practice*. 2009;10:52.
43. van Weel-Baumgarten EM, van den Bosch WJ, van den Hoogen HJ, Zitman FG. The validity of the diagnosis of depression in general practice: is using criteria for diagnosis as a routine the answer? *Br J Gen Pract*. 2000;50:284-7.
44. Kessler D, Lloyd K, Lewis G, Gray DP. Cross sectional study of symptom attribution and recognition of depression and anxiety in primary care. *BMJ*. 1999;318:436-9.
45. Wittink MN, Barg FK, Gallo JJ. Unwritten rules of talking to doctors about depression: integrating qualitative and quantitative methods. *Ann Fam Med*. 2006;4:302-9.
46. Menear M, Dore I, Cloutier AM, Perrier L, Roberge P, Duhoux A, et al. Chronic physical comorbidity burden and the quality of depression treatment in primary care: a systematic review. *J Psychosom Res*. 2015;78:314-23.

47. Barkow K, Heun R, Ustun TB, Berger M, Bermejo I, Gaebel W, et al. Identification of somatic and anxiety symptoms which contribute to the detection of depression in primary health care. *Eur Psychiatry*. 2004;19:250-7.
48. Hutton C, Gunn J. Do longer consultations improve the management of psychological problems in general practice? A systematic literature review. *BMC Health Serv Res*. 2007;7:71.
49. Pollock K, Grime J. GPs' perspectives on managing time in consultations with patients suffering from depression: a qualitative study. *Family practice*. 2003;20:262-9.
50. Pollock K, Grime J. Patients' perceptions of entitlement to time in general practice consultations for depression: qualitative study. *BMJ*. 2002;325:687.
51. Balestrieri M, Baldacci S, Bellomo A, Bellantuono C, Conti L, Perugi G, et al. Clinical vs. structured interview on anxiety and affective disorders by primary care physicians. understanding diagnostic discordance. *Epidemiologia e psichiatria sociale*. 2007;16:144-51.
52. Berardi D, Menchetti M, Cevenini N, Scaini S, Versari M, De Ronchi D. Increased recognition of depression in primary care. Comparison between primary-care physician and ICD-10 diagnosis of depression. *Psychother Psychosom*. 2005;74:225-30.
53. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet*. 2009;374:609-19.
54. Castro-Rodriguez JI, Olariu E, Garnier-Lacueva C, Martin-Lopez LM, Perez-Sola V, Alonso J, et al. Diagnostic accuracy and adequacy of treatment of depressive and anxiety disorders: A comparison of primary care and specialized care patients. *J Affect Disorders*. 2015;172:462-71.
55. Olariu E, Forero CG, Castro-Rodriguez JI, Rodrigo-Calvo MT, Alvarez P, Martin-Lopez LM, et al. Detection of Anxiety Disorders in Primary Care: A Meta-Analysis of Assisted and Unassisted Diagnoses. *Depress Anxiety*. 2015;32:471-84.
56. Simon GE, Goldberg D, Tiemens BG, Ustun TB. Outcomes of recognized and unrecognized depression in an international primary care study. *Gen Hospital Psychiatry*. 1999;21:97-105.
57. Solberg LI, Korsen N, Oxman TE, Fischer LR, Bartels S. The need for a system in the care of depression. *J Fam Practice*. 1999;48:973-9.
58. Garzonis K, Mann E, Wyrzykowska A, Kanellakis P. Improving Patient Outcomes: Effectively Training Healthcare Staff in Psychological Practice Skills: A Mixed Systematic Literature Review. *Eur J Psychol*. 2015;11:535-56.
59. Gilbody S, House AO, Sheldon TA. Screening and case finding instruments for depression. *Cochrane Database Syst Rev*. 2005(4):CD002792.
60. Spiegel W, Tonies H, Scherer M, Katschnig H. Learning by doing: a novel approach to improving general practitioners' diagnostic skills for common mental disorders. *Wiener klinische Wochenschrift*. 2007;119:117-23.
61. Kamphuis MH, Stegenga BT, Zuithoff NP, King M, Nazareth I, de Wit NJ, et al. Does recognition of depression in primary care affect outcome? The PREDICT-NL study. *Family practice*. 2012;29:16-23.
62. Smolders M, Laurant M, Verhaak P, Prins M, van Marwijk H, Penninx B, et al. Adherence to evidence-based guidelines for depression and anxiety disorders is associated with recording of the diagnosis. *Gen Hosp Psychiatry*. 2009;31:460-9.

63. Spitzer RL, Williams JB, Kroenke K, Linzer M, deGruy FV, 3rd, Hahn SR, et al. Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. *JAMA*. 1994;272:1749-56.
64. Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression. Two questions are as good as many. *J Gen Intern Med*. 1997;12:439-45.
65. Beck AT, Steer RA, Ball R, Ranieri W. Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *J Pers Assess*. 1996;67:588-97.
66. Gilbody S, Sheldon T, House A. Screening and case-finding instruments for depression: a meta-analysis. *CMAJ*. 2008;178:997-1003.
67. O'Connor EA, Whitlock EP, Beil TL, Gaynes BN. Screening for depression in adult patients in primary care settings: a systematic evidence review. *Ann Intern Med*. 2009;151:793-803.
68. Silverstone PH, Rittenbach K, Suen VY, Moretzsohn A, Cribben I, Bercov M, et al. Depression Outcomes in Adults Attending Family Practice Were Not Improved by Screening, Stepped-Care, or Online CBT during a 12-Week Study when Compared to Controls in a Randomized Trial. *Front Psychiatry*. 2017;8:32.
69. Kashner TM, Rush AJ, Suris A, Biggs MM, Gajewski VL, Hooker DJ, et al. Impact of structured clinical interviews on physicians' practices in community mental health settings. *Psychiatr Serv*. 2003;54:712-8.
70. Swedish Council on Health Technology Assessment (SBU). [Diagnosis and follow-up of affective disorders -a systematic review]. Stockholm: SBU report 212; 2012.
71. Spitzer RL. Psychiatric diagnosis: are clinicians still necessary? *Comprehens Psychiatry*. 1983;24:399-411.
72. Leckman JF, Sholomskas D, Thompson WD, Belanger A, Weissman MM. Best estimate of lifetime psychiatric diagnosis: a methodological study. *Arch Gen Psychiatry*. 1982;39:879-83.
73. Agency for health Care Policy and Research Depression Guideline Panel. Depression in primary care: Vol 1. Detection and diagnosis (Clinical Practice Guideline No 5; ACHPR Publication No. 93 0550). Rockville, MD: U.S. Department of health and human services, Public Health Service; 1993.
74. British Medical Association. NHS Employers. Revisions to the GMS contract 2006/2007. British medical Association, London 2006.
75. National Board for Health and Welfare. National guidelines for care of depression and anxiety disorders [in Swedish]. National board for health and welfare; Västerås 2010.
76. National Board for Health and Welfare. [Antidepressiva läkemedel vid psykisk ohälsa. Studier av praxis i primärvården]. 2006.
77. Hassler E and Krakau I. [Primärvården, psykiatrin och den deprimerade patienten – en studie av vårdpraxis i ett avgränsat geografiskt område]. [unpublished work].
78. Dowrick C, Leydon GM, McBride A, Howe A, Burgess H, Clarke P, et al. Patients' and doctors' views on depression severity questionnaires incentivised in UK quality and outcomes framework: qualitative study. *BMJ*. 2009;338:b663.

79. Leydon GM, Dowrick CF, McBride AS, Burgess HJ, Howe AC, Clarke PD, et al. Questionnaire severity measures for depression: a threat to the doctor-patient relationship? *Br J Gen Pract.* 2011;61:117-23.
80. Mitchell C, Dwyer R, Hagan T, Mathers N. Impact of the QOF and the NICE guideline in the diagnosis and management of depression: a qualitative study. *Br J Gen Pract.* 2011;61:e279-89.
81. Maxwell M, Harris F, Hibberd C, Donaghy E, Pratt R, Williams C, et al. A qualitative study of primary care professionals' views of case finding for depression in patients with diabetes or coronary heart disease in the UK. *BMC family practice.* 2013;14:46.
82. Wittkamp KA, van Zwieten M, Smits FT, Schene AH, Huyser J, van Weert HC. Patients' view on screening for depression in general practice. *Family practice.* 2008;25:438-44.
83. Wood F, Pill R, Prior L, Lewis G. Patients' opinions of the use of psychiatric case-finding questionnaires in general practice. *Health expectations.* 2002;5:282-8.
84. Wikberg C, Pettersson A, Westman J, Bjorkelund C, Petersson EL. Patients' perspectives on the use of the Montgomery-Asberg depression rating scale self-assessment version in primary care. *Scand J Prim Health Care.* 2016;34:434-42.
85. Tavabie JA, Tavabie OD. Improving care in depression: qualitative study investigating the effects of using a mental health questionnaire. *Qual Prim Care.* 2009;17:251-61.
86. Eccles MP, Armstrong D, Baker R, Cleary K, Davies H, Davies S, et al. An implementation research agenda. *Implement Sci.* 2009;4:18.
87. Nilsen P. Making sense of implementation theories, models and frameworks. *Implement Sci.* 2015;10:53.
88. Swedish Council on Health Technology Assessment (SBU). Implementation support for psychiatric evidence in primary care [in Swedish]. Stockholm: SBU; 2012.
89. Prestwich A, Sniehotta FF, Whittington C, Dombrowski SU, Rogers L, Michie S. Does theory influence the effectiveness of health behavior interventions? Meta-analysis. *Health Psychol.* 2014;33:465-74.
90. Birken SA, Powell BJ, Pesseau J, Kirk MA, Lorencatto F, Gould NJ, et al. Combined use of the Consolidated Framework for Implementation Research (CFIR) and the Theoretical Domains Framework (TDF): a systematic review. *Implement Sci.* 2017;12:2.
91. Flottorp SA, Oxman AD, Krause J, Musila NR, Wensing M, Godycki-Cwirko M, et al. A checklist for identifying determinants of practice: a systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. *Implement Sci.* 2013;8:35.
92. Gibson O, Lisy K, Davy C, Aromataris E, Kite E, Lockwood C, et al. Enablers and barriers to the implementation of primary health care interventions for Indigenous people with chronic diseases: a systematic review. *Implement Sci.* 2015;10:71.
93. May CR, Johnson M, Finch T. Implementation, context and complexity. *Implement Sci.* 2016;11:141.
94. Baker R, Camosso-Stefinovic J, Gillies C, Shaw EJ, Cheater F, Flottorp S, et al. Tailored interventions to overcome identified barriers to change: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev.* 2010(3):CD005470.

95. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q.* 2004;82:581-629.
96. Moullin JC, Sabater-Hernandez D, Fernandez-Llimos F, Benrimoj SI. A systematic review of implementation frameworks of innovations in healthcare and resulting generic implementation framework. *Health research policy and systems / BioMed Central.* 2015;13:16.
97. Kitson A, Harvey G, McCormack B. Enabling the implementation of evidence based practice: a conceptual framework. *Qual Health Care.* 1998;7:149-58.
98. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci.* 2009;4:50.
99. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. *Implement Sci.* 2012;7:37.
100. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci.* 2011;6:42.
101. National Board for Health and Welfare. National guidelines for schizophrenia and schizophrenia-type conditions. [In Swedish]. Edita Västra Aros, Västerås 2011.
102. Sandstrom B, Willman A, Svensson B, Borglin G. 'How do we know if this is the best?' Mental health-care professionals' views on national guidelines for psychosocial interventions. *Int J Ment Health Nurs.* 2014;23:221-31.
103. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med.* 2013;46:81-95.
104. Michie S, Atkins L, West R. *The Behaviour Change Wheel- a guide to designing interventions.* Great Britain: Silverback publishing; 2014.
105. Wensing M, Bosch M, Grol R. Developing and selecting interventions for translating knowledge to action. *CMAJ.* 2010;182:E85-8.
106. Sinnema H, Majo MC, Volker D, Hoogendoorn A, Terluin B, Wensing M, et al. Effectiveness of a tailored implementation programme to improve recognition, diagnosis and treatment of anxiety and depression in general practice: a cluster randomised controlled trial. *Implement Sci.* 2015;10:33.
107. Shirazi M, Lonka K, Parikh SV, Ristner G, Alaeddini F, Sadeghi M, et al. A tailored educational intervention improves doctor's performance in managing depression: a randomized controlled trial. *J Eval Clin Pract.* 2013;19:16-24.
108. Baker R, Reddish S, Robertson N, Hearnshaw H, Jones B. Randomised controlled trial of tailored strategies to implement guidelines for the management of patients with depression in general practice. *Br J Gen Pract.* 2001;51:737-41.
109. Baker R, Camosso-Stefinovic J, Gillies C, Shaw EJ, Cheater F, Flottorp S, et al. Tailored interventions to address determinants of practice. *Cochrane Database Syst Rev.* 2015(4):CD005470.
110. Wensing M. The Tailored Implementation in Chronic Diseases (TICD) project: introduction and main findings. *Implement Sci.* 2017;12:5.

111. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*. 2015;350:h1258.
112. Rogers EM. *Diffusion of Innovations*. 5th edition. New York: Free Press; 1993.
113. Ajzen I. Nature and operation of attitudes. *Annu Rev Psychol*. 2001;52:27-58.
114. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev*. 1977;84:191-215.
115. Nilsen P, Roback K, Brostrom A, Ellstrom PE. Creatures of habit: accounting for the role of habit in implementation research on clinical behaviour change. *Implement Sci*. 2012;7:53.
116. Wensing M, Wollersheim H, Grol R. Organizational interventions to implement improvements in patient care: a structured review of reviews. *Implement Sci*. 2006;1:2.
117. Chauhan BF, Jeyaraman M, Mann AS, Lys J, Skidmore B, Sibley KM, et al. Behavior change interventions and policies influencing primary healthcare professionals' practice-an overview of reviews. *Implement Sci*. 2017;12:3.
118. Grol R, Grimshaw J. Evidence-based implementation of evidence-based medicine. *J Comm J Qual Improv*. 1999;25:503-13.
119. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet*. 2003;362:1225-30.
120. Mickan SB, A.; Glasziou, P. Patterns of 'leakage' in the utilisation of clinical guidelines: a systematic review. *Postgrad Med J*. 2011;87:670-9.
121. Lau R, Stevenson F, Ong BN, Dziedzic K, Treweek S, Eldridge S, et al. Achieving change in primary care--causes of the evidence to practice gap: systematic reviews of reviews. *Implement Sci*. 2016;11:40.
122. Sinnema H, Terluin B, Wensing M, Volker D, Franx G, van Balkom A, et al. Systematic tailoring for the implementation of guideline recommendations for anxiety and depressive disorders in general practice: perceived usefulness of tailored interventions. *BMC family practice*. 2013;14:94.
123. Lugtenberg M, Zegers-van Schaick JM, Westert GP, Burgers JS. Why don't physicians adhere to guideline recommendations in practice? An analysis of barriers among Dutch general practitioners. *Implement Sci*. 2009;4:54.
124. Zwolsman S, te Pas E, Hooft L, Wieringa-de Waard M, van Dijk N. Barriers to GPs' use of evidence-based medicine: a systematic review. *Br J Gen Pract*. 2012;62:e511-21.
125. Toner RS, C.; Acton, S.; Blenkiron, B. Do general practitioners adhere to NICE guidelines for depression? Systematic Questionnaire Survey. *Prim Health Care Res Dev*. 2010;11:123-31.
126. Gunn JM, Palmer VJ, Dowrick CF, Herrman HE, Griffiths FE, Kokanovic R, et al. Embedding effective depression care: using theory for primary care organisational and systems change. *Implement Sci*. 2010;5:62.
127. May CR, Mair F, Finch T, MacFarlane A, Dowrick C, Treweek S, et al. Development of a theory of implementation and integration: Normalization Process Theory. *Implement Sci*. 2009;4:29.

128. Denscombe M. A Research Paradigm for the Mixed Methods Approach. *J Mixed Methods Research*. 2008;2:270-83.
129. Greene JC. Toward a Conceptual Framework for Mixed-Method Evaluation Designs. *Educational Evaluation and Policy Analysis*. 1989;11:255-74.
130. Guest G. Describing Mixed Methods Research: An Alternative to Typologies. *J Mixed Methods Research*. 2012;7:141-51.
131. Tariq SaW, J. Using mixed methods in health research. *J R Soc Med Short Reports*. 2013.
132. Creswell JWK, A.C. ;Plano Clark, V.L.; Clegg Smith, K. Best practices for mixed methods research in the health sciences. Research Office of Behavioural and Social Sciences Research. NIH.
133. Dures E, Rumsey N, Morris M, Gleeson K. Mixed methods in health psychology: theoretical and practical considerations of the third paradigm. *J Health Psychol*. 2011;16:332-41.
134. Bishop FL. Using mixed methods research designs in health psychology: an illustrated discussion from a pragmatist perspective. *Br J Health Psychol*. 2015;20:5-20.
135. O'Cathain A, Murphy E, Nicholl J. Three techniques for integrating data in mixed methods studies. *BMJ*. 2010;341:c4587.
136. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336:924-6.
137. Patton MQ. *Qualitative Research & Evaluation Methods*. 3rd edition. Thousand Oaks, California: Sage Publications; 2002.
138. Lincoln YS, Guba E.G But Is It rigorous? Trustworthiness and Authenticity in Naturalistic Evaluation. *New Directions for Program Evaluation*. 1986:73-84.
139. Schunemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ*. 2008;336:1106-10.
140. Swedish Council on Health Technology Assessment (SBU). Treatment of Depression [In Swedish]. Stockholm: SBU report 166; 2004.
141. Glasziou P, Chalmers I, Altman DG, Bastian H, Boutron I, Brice A, et al. Taking healthcare interventions from trial to practice. *BMJ*. 2010;341:c3852.
142. Malterud K. Systematic text condensation: a strategy for qualitative analysis. *Scand J Public health*. 2012;40:795-805.
143. Ministry of Health and Social Affairs. The Patient Law (2014:821) [in Swedish]. *Svensk författningssamling* 2014:82
144. Ministry of Health and Social Affairs. [Lag om valfrihetssystem]. *Svensk författningssamling* 2008:962 .
145. Medicinska Programrådet. Regional guideline for depression and bipolar disorders [in Swedish]. Stockholm County Council. 2007.

146. Medicinska Programrådet. Regional guidelines for anxiety disorders [in Swedish]. Stockholm County Council. 2011.
147. Wikberg C, Westman J, Petersson EL, Larsson ME, Andre M, Eggertsen R, et al. Use of a self-rating scale to monitor depression severity in recurrent GP consultations in primary care - does it really make a difference? A randomised controlled study. *BMC family practice*. 2017;18:6.
148. Sundquist K, Malmstrom M, Johansson SE, Sundquist J. Care Need Index, a useful tool for the distribution of primary health care resources. *J Epidemiol Community health*. 2003;57:347-52.
149. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 1998;59 Suppl 20:22-33.
150. de Azevedo Marques JM, Zuardi AW. Validity and applicability of the Mini International Neuropsychiatric Interview administered by family medicine residents in primary health care in Brazil. *Gen Hosp Psychiatry*. 2008;30:303-10.
151. Otsubo T, Tanaka K, Koda R, Shinoda J, Sano N, Tanaka S, et al. Reliability and validity of Japanese version of the Mini-International Neuropsychiatric Interview. *Psychiatry Clin Neurosci*. 2005;59:517-26.
152. Baas KD, Wittkamp KA, van Weert HC, Lucassen P, Huyser J, van den Hoogen H, et al. Screening for depression in high-risk groups: prospective cohort study in general practice. *Br J Psychiatry*. 2009;194:399-403.
153. Gill D, Sharpe M. Frequent consulters in general practice: a systematic review of studies of prevalence, associations and outcome. *J Psychosom Res*. 1999;47:115-30.
154. Gili M, Luciano JV, Serrano MJ, Jimenez R, Bauza N, Roca M. Mental disorders among frequent attenders in primary care: a comparison with routine attenders. *J Nerv Ment Dis*. 2011;199:744-9.
155. Karlsson H, Lehtinen V, Joukamaa M. Psychiatric morbidity among frequent attender patients in primary care. *Gen Hosp Psychiatry*. 1995;17:19-25.
156. Suppiger A, In-Albon T, Hendriksen S, Hermann E, Margraf J, Schneider S. Acceptance of structured diagnostic interviews for mental disorders in clinical practice and research settings. *Behavior therapy*. 2009;40:272-9.
157. Kitzinger J. Qualitative research. Introducing focus groups. *BMJ*. 1995;311:299-302.
158. Krueger R, Casey M. *Focus Groups: A Practical Guide for Applied Research*. Los Angeles: SAGE Publications Inc; 2009.
159. Wibeck V. *Fokusgrupper: Om fokuserade gruppintervjuer som undersökningsmetod*. Lund: Studentlitteratur; 2010.
160. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. 2009;339:b2535.
161. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ*. 2003;326:41-4.

162. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol.* 2003;3:25.
163. Gilbody S, Sheldon T, Wessely S. Should we screen for depression? *BMJ.* 2006;332:1027-30.
164. Deeks JJ. Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests. *BMJ.* 2001;323:157-62.
165. Macaskill PG, CA; Deeks, JJ; Harbord, RM; Takwoingi, Y. Chapter 10. Analysing and presenting results. *Cochrane Handbook for systematic reviews of diagnostic test accuracy: Cochrane Collaboration; 2010.*
166. Harbord RM, Whiting P, Sterne JA, Egger M, Deeks JJ, Shang A, et al. An empirical comparison of methods for meta-analysis of diagnostic accuracy showed hierarchical models are necessary. *J Clin Epidemiol.* 2008;61:1095-103.
167. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol.* 2005;58:982-90.
168. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med.* 2001;20:2865-84.
169. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics.* 2007;8:239-51.
170. Takwoingi Y Deeks JJ. METADAS: A SAS macro for meta-analysis of diagnostic accuracy studies. <http://srdta.cochrane.org/software-development>. Cochrane Collaboration 2008.
171. Zamora J, Abaira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol.* 2006;6:31.
172. Giorgi A. Sketch of a psychological phenomenological method. Pittsburgh, PA: Duquesne University Press; 1985.
173. Elo S, Kyngas H. The qualitative content analysis process. *J Adv Nursing.* 2008;62:107-15.
174. Krippendorff K. Content analysis: An introduction to its methodology. 3rd edition. Thousand Oaks, CA: SAGE Publications; 2013.
175. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res.* 2005;15:1277-88.
176. Farmer T, Robinson K, Elliott SJ, Eyles J. Developing and implementing a triangulation protocol for qualitative health research. *Qual Health Res.* 2006;16:377-94.
177. Rickham PP. Human Experimentation. Code of Ethics of the World Medical Association. Declaration of Helsinki. *Br Med J.* 1964;2:177.
178. Spitzer RL, Williams JB, Gibbon M, First MB. The Structured Clinical Interview for DSM-III-R (SCID). I: History, rationale, and description. *Arch Gen Psychiatry.* 1992;49:624-9.

179. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA. 1999;282:1737-44.
180. Pettersson A, Bostrom KB, Gustavsson P, Ekselius L. Which instruments to support diagnosis of depression have sufficient accuracy? A systematic review. Nord J Psychiatry. 2015;69:497-508.
181. Baik SY, Gonzales JJ, Bowers BJ, Anthony JS, Tidjani B, Susman JL. Reinvention of depression instruments by primary care clinicians. Ann Fam Med. 2010;8:224-30.
182. Amorim P, Lecrubier Y, Weiller E, Hergueta T, Sheehan D. DSM-III-R Psychotic Disorders: procedural validity of the Mini International Neuropsychiatric Interview (MINI). Concordance and causes for discordance with the CIDI. Eur Psychiatry. 1998;13:26-34.
183. Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. CMAJ. 2012;184:E191-6.
184. Meader N, Mitchell AJ, Chew-Graham C, Goldberg D, Rizzo M, Bird V, et al. Case identification of depression in patients with chronic physical health problems: a diagnostic accuracy meta-analysis of 113 studies. Br J Gen Pract. 2011;61:e808-20.
185. Mitchell AJ, Yadegarfar M, Gill J, Stubbs B. Case finding and screening clinical utility of the Patient Health Questionnaire (PHQ-9 and PHQ-2) for depression in primary care: a diagnostic meta-analysis of 40 studies. BJ Psych Open. 2016;2:127-38.
186. Moriarty AS, Gilbody S, McMillan D, Manea L. Screening and case finding for major depressive disorder using the Patient Health Questionnaire (PHQ-9): a meta-analysis. Gen Hosp Psychiatry. 2015;37:567-76.
187. Mordal J, Gundersen O, Bramness JG. Norwegian version of the Mini-International Neuropsychiatric Interview: feasibility, acceptability and test-retest reliability in an acute psychiatric ward. Eur Psychiatry. 2010;25:172-7.
188. Pinninti NR, Madison H, Musser E, Rissmiller D. MINI International Neuropsychiatric Schedule: clinical utility and patient acceptance. Eur Psychiatry. 2003;18:361-4.
189. Zimmerman M, Mattia JI. Psychiatric diagnosis in clinical practice: is comorbidity being missed? Compr Psychiatry. 1999;40:182-91.
190. Ramirez Basco M, Bostic JQ, Davies D, Rush AJ, Witte B, Hendrickse W, et al. Methods to improve diagnostic accuracy in a community mental health setting. Am J Psychiatry. 2000;157:1599-605.
191. Miller PR, Dasher R, Collins R, Griffiths P, Brown F. Inpatient diagnostic assessments: 1. Accuracy of structured vs. unstructured interviews. Psychiatry Res. 2001;105:255-64.
192. Rettew DC, Lynch AD, Achenbach TM, Dumenci L, Ivanova MY. Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. Int J Methods Psychiatr Res. 2009;18:169-84.
193. Tiemens BG, VonKorff M, Lin EH. Diagnosis of depression by primary care physicians versus a structured diagnostic interview. Understanding discordance. Gen Hosp Psychiatry. 1999;21:87-96.
194. Bakker IM, Terluin B, van Marwijk HW, van Mechelen W, Stalman WA. Test-retest reliability of the PRIME-MD: limitations in diagnosing mental disorders in primary care. Eur J Public Health. 2009;19:303-7.

195. Jonasson B, Jonasson U, Ekselius L, von Knorring L. The feasibility of a new intake routine to assess substance use disorders by means of a structured interview. *Gen Hosp Psychiatry*. 1997;19:36-41.
196. Marshall R, Spitzer R, Vaughan S, Vaughan R, Mellman L, MacKinnon R, et al. Assessing the subjective experience of being a participant in psychiatric research. *Am J Psychiatry*. 2001;158:319-21.
197. Sorensen MJ, Thomsen PH, Bilenberg N. Parent and child acceptability and staff evaluation of K-SADS-PL: a pilot study. *Eur child & adolesc Psychiatry*. 2007;16:293-7.
198. Sharma VK, Lepping P, Cummins AG, Copeland JR, Parhee R, Mottram P. The Global Mental Health Assessment Tool--Primary Care Version (GMHAT/PC). Development, reliability and validity. *World Psychiatry*. 2004;3:115-9.
199. Sharma VK, Lepping P, Krishna M, Durrani S, Copeland JR, Mottram P, et al. Mental health diagnosis by nurses using the Global Mental Health Assessment Tool: a validity and feasibility study. *Br J Gen Pract*. 2008;58:411-6.
200. Kitson AL, Harvey G. Methods to Succeed in Effective Knowledge Translation in Clinical Practice. *J Nurs Scholarsh*. 2016;48:294-302.
201. Lesser I. Cultural considerations using the Structured Clinical Interview for DSM-III for mood and anxiety disorders assessment. *J Psychopathol Behav Assess*. 1997;19:149-60.
202. Baas KD, Cramer AO, Koeter MW, van de Lisdonk EH, van Weert HC, Schene AH. Measurement invariance with respect to ethnicity of the Patient Health Questionnaire-9 (PHQ-9). *J Affect Disorders*. 2011;129:229-35.
203. Maters GA, Sanderman R, Kim AY, Coyne JC. Problems in cross-cultural use of the hospital anxiety and depression scale: "no butterflies in the desert". *PLoS One*. 2013;8(8):e70975.
204. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care*. 2007;19:349-57.
205. Fugard AJ, Potts, H. Supporting thinking on sample sizes for thematic analyses: a quantitative tool. *Int J Soc Res Methodol*. 2015;18:669-84.
206. Malterud K, Siersma VD, Guassora AD. Sample Size in Qualitative Interview Studies: Guided by Information Power. *Qual Health Res*. 2015 Nov 27. pii: 1049732315617444.
207. Colquhoun HL, Squires JE, Kolehmainen N, Fraser C, Grimshaw JM. Methods for designing interventions to change healthcare professionals' behaviour: a systematic review. *Implement Sci*. 2017;12:30.
208. Harris M, Lawn SJ, Morello A, Battersby MW, Ratcliffe J, McEvoy RD, et al. Practice change in chronic conditions care: an appraisal of theories. *BMC Health Serv Res*. 2017;17:170.
209. National Board for Health and Welfare. National Guidelines for care of depression and anxiety disorders [in Swedish]. Västerås. 2016.
210. National Board for Health and Welfare. Use of instruments for depression in primary care - a web based survey [In Swedish, unpublished work]. 2016.

211. Fredriksson M, Eldh AC, Vengberg S, Dahlstrom T, Halford C, Wallin L, et al. Local politico-administrative perspectives on quality improvement based on national registry data in Sweden: a qualitative study using the Consolidated Framework for Implementation Research. *Implement Sci.* 2014;9:189.