Depression and anxiety coexisting with osteoarthritis in primary care: from recognition to management

Magdalena Rzewuska

A thesis submitted for the degree of

Doctor of Philosophy

June 2013

Arthritis Research UK Primary Care Centre

Keele University

Declaration

This PhD was funded by a Keele University Acorn fellowship (Studentship Reference: 2008-13) obtained at the Arthritis Research UK Primary Care Centre by Professor Christian Mallen and colleagues. The original topic of this PhD project was titled *Depression and joint pain in primary care* and formed part of the submission for this studentship.

Throughout the PhD project, with guidance from my supervisors Professor Christian Mallen, Professor George Peat, Dr John Belcher and statistical advisor Dr Vicky Strauss, I developed the ideas around, and managed the direction of the thesis.

The PROG-RES data was collected prior to my involvement in the study. I had no control over the content of the surveys. In this thesis I have approached the use of the PROG-RES data as a secondary analysis of previously collected data. Data was provided by the data custodian Dr Sara Muller.

My supervisors advised me on the planning of all analyses and on the writing and presentation of chapters. I conducted all analyses and wrote the chapters myself. I received guidance on the search strategy for the literature review from my supervisors and a specialist librarian Dr Rachel Gick. Professor Christian Mallen, Professor George Peat, Dr Barbara Nicholl and Dr Kay Benyon acted as independent reviewers. Dr Ross Harris provided advice by email on generating forest plots of prevalence estimates. Dr Vicky Strauss provided access to the Latent GOLD software and expert advice on data analyses. Professor Christian Mallen offered his expert clinical advice on the selection of LCGA cluster solution and guidance on the list of Read codes and clinical search terms for the medial record review.

Abstract

Osteoarthritis (OA), depression and anxiety are common problems in primary care. OA coexisting with depressive and/or anxiety symptoms has detrimental consequences to the individual. To inform recognition and management of these important problems in primary care a better understanding of their coexistence is needed.

A systematic review with meta-analysis was undertaken to determine the prevalence of depression and anxiety in adults with OA/joint pain in the community. Elevated anxiety symptoms were more common (45%) than depression symptoms (24%) in persons with OA/joint pain. Sources of between-study variance include methods of ascertainment and geographical location. A review of measurement properties of several recommended patient-reported depression and anxiety measures found evidence to support properties in some populations, but some critical properties warrant investigation in adults with OA in the community.

A secondary data analysis was conducted for older consecutive primary care patients with musculoskeletal pain recruited to a cohort (n=443) of the PROGnostic Research study. Latent Class Growth Analyses identified clusters of individuals who exhibited different trajectories of anxiety and depression symptoms over a 12-month period: three anxiety and two depression symptom trajectories. In total, 56% and 63% of participants experienced persistent anxiety and depression symptoms respectively for at least 12 months. Pain characteristics and coping strategies were the most prominent risk factors for persistent anxiety and depression symptoms. With the aim of identifying individuals with sub-

threshold persistent anxiety and depression symptoms, characteristics predisposing to symptoms persistence may be considered.

A medical records review found that only half of all older musculoskeletal patients with persistent anxiety and depression symptoms have their mental health problems detected by their GP. Frequent consulters and those with more severe anxiety were more likely to be detected. This reinforces the need to recognise and manage OA coexisting with depression and/or anxiety by patients and health professionals alike.

Acknowledgements

I would like to thank my supervisors Professor Christian Mallen, Professor George Peat and Dr John Belcher for their patience, invaluable support and guidance on the drafts of this thesis. I am particularly grateful to my lead academic supervisor Professor Christian Mallen for giving me the opportunity to complete this PhD and enabling my growth as an independent researcher. I would also like to thank my statistical advisor Dr Vicky Strauss and PROG-RES data custodian Dr Sara Muller. I am very grateful to Dr Barbara Nicholl and Dr Kay Benyon for acting as independent reviewers in the systematic review and for random acts of compassion and support. I am very grateful to patients, GP practices, informatics and support staff involved in the PROG-RES study. Thanks to all the staff and students at Arthritis Research UK Primary Care Centre for their encouragement and support.

Finally, I would like to express my thanks to Tulinho, all my friends (particularly Gabriela and Nica) and my family, for their understanding, love and care. I am dedicating this thesis to my personal inspiration Marie Skłodowska-Curie, and to all other fellow Pols who left their homeland in search for meaning and fulfilment.

Context of the thesis

At the age of 18 I took on a role of a child supporter and family counsellor. This led me to undertaking BA in Counseling at the University of Szczecin in Poland, graduating in 2004. Before I entered into clinical practice, I decided to broaden my psychological knowledge, and thus I undertook BSc in Neuropsychology at the University of Central Lancashire, graduating in 2006. During this period I worked as an activities coordinator at a care home specialising in older people with mental health problems. Through my work with older people, I developed strong research interests in physiological and psychological responses to stress. This led me to conduct a thesis project in this area - later shortlisted by the British Psychological Society as one of the 10 best neuropsychology projects in the UK. Encouraged by my early academic successes I decided to make a research contribution to an improved mental health care for older people. I successfully applied for a three-year full-time PhD studentship advertised at the Arthritis Research UK Primary Care Centre. My biggest challenges were taking a secondary data approach and adopting a primary care view on mental health well-being, with a strong emphasis on implications for clinical practice.

Table of contents

List of Tables

Chapter two: Prevalence of depression and anxiety in adults with osteoarthritis/joint pain: a systematic review and meta-analyses

Chapter three: Self-report measures for depression and anxiety symptom screening and assessment in osteoarthritis: a narrative review of selected measurement properties

Chapter four: The course of anxiety and depression symptoms in older patients presenting to general practice with musculoskeletal pain
Part 1: Source and suitability of the PROG-RES dataset

Chapter five: The course of anxiety and depression symptoms in older patients presenting to general practice with musculoskeletal pain
Part 2: A latent class growth analysis

Chapter six: The course of anxiety and depression symptoms in older patients presenting to general practice with musculoskeletal pain
Part 3: Factors associated with the course of anxiety and depression symptoms

Chapter seven: Documented detection of depression and anxiety in older adults consulting with musculoskeletal pain: analyses of medical record data

Chapter eight: Discussion, conclusions and recommendations

List of Figures

Chapter one: Thesis outline, key concepts and operational frameworks

Chapter two: Prevalence of depression and anxiety in adults with osteoarthritis/joint pain: a systematic review and meta-analyses

Chapter four: The course of anxiety and depression symptoms in older patients presenting to general practice with musculoskeletal pain

Part 1: Source and suitability of the PROG-RES dataset

Chapter five: The course of anxiety and depression symptoms in older patients presenting to general practice with musculoskeletal pain

Part 2: A latent class growth analysis

Chapter seven: Documented detection of depression and anxiety in older adults consulting with musculoskeletal pain: analyses of medical record data

Chapter eight: Discussion, conclusions and recommendations

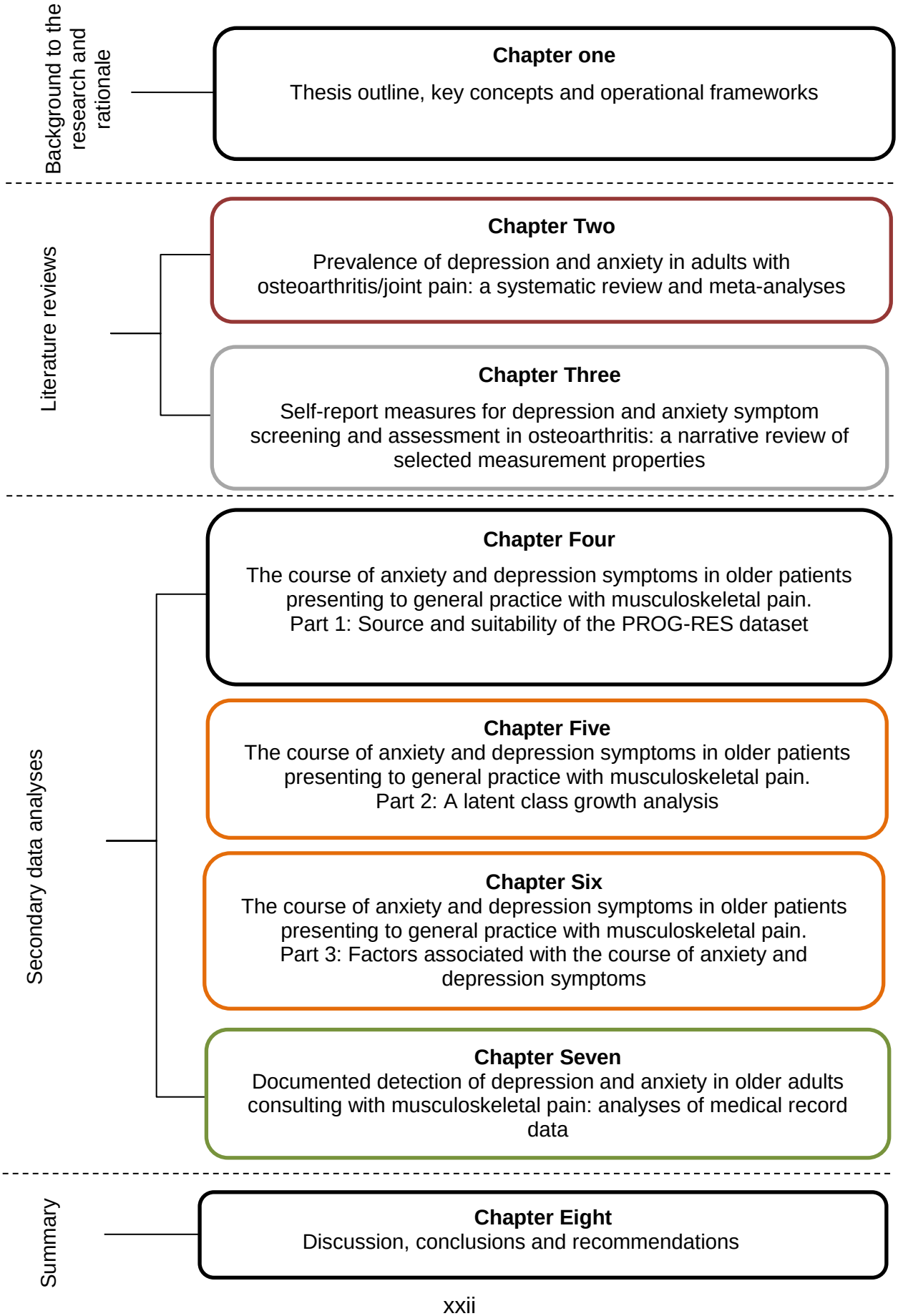
List of Abbreviations

ACR	American College of Rheumatology
AIMS (D/A)	Arthritis Impact Measurement Scales (depression/anxiety)
APA	American Psychiatric Association
ASD	Acute stress disorder
AUDADIS	Alcohol Use Disorders and Associated Disabilities Interview
BDI (PC)	Beck depression inventory (Primary care)
BIC	Bayesian Information Criterion
BL	Baseline
BLRT	Bootstrap likelihood ratio test
BN	Barbara Nicholl
BNF	British National Formulary
CALM	Coordinated Anxiety Learning and Management
CB-SCIDI	Patient version of Chinese-bilingual Clinical Interview schedule
CESD	Center for Epidemiologic Studies Depression Scale
CI	Confidence Interval
(C/M) CIDI	(Computer/Munich) Composite International Diagnostic Interview
CIS	Semi-structured clinical interview schedule
CM	Christian Mallen
CPG	Chronic Pain Grade
CSQ	Coping Strategies Questionnaire
DSM	Diagnostic and Statistical Manual of Mental Disorders
EM	Expectation-maximization
EMR	Electronic medical records
ES	Estimate
4DSQ (A/D)	Four dimensional Symptoms Questionnaire (anxiety/depression)
FUP	Follow-up
GAD	Generalised anxiety disorder
GADS (A/D)	Goldberg Anxiety and Depression Scale (anxiety/depression)
GDS	Geriatric Depression Scale

GMM	Growth mixture modelling
GMSA	Short Geriatric Mental State Examination
GMP	George Peat
GP	General Practitioner
HADS	Hospital Anxiety and Depression Scale
HAM-A	Hamilton Anxiety rating scale
IAPT	Improving Access to Psychological Therapies
ICC	Intra-class correlation
ICD	International Classification of Diseases
IMPACT	Improving Mood-Promoting Access to Collaborative Treatment
IQR	Inter-quartile Range
IRGL (A/D)	Influence of Rheumatic Diseases on Health and Lifestyle (anxiety/depression)
KB	Kay Benyon
LCGA	Latent class growth analyses
LL	Log-likelihood
LLCA	Longitudinal latent class analyses
LR	Likelihood ratio
MD	Major depression
MED	Median difference
MINI	Mini -International Neuropsychiatric Interview
ML	Maximum likelihood
MOOSE	Meta-analysis Of Observational Studies in Epidemiology
MR	Magdalena Rzewuska
MSK	Musculoskeletal complaints
N/A	Not applicable
NCC-CC	National Collaborating Centre for Chronic Conditions
NCC-MH	National Collaborating Centre for Mental Health
NHS	National Health Service
NICE	National Institute of Health and Clinical Excellence
NOS	Not otherwise specified
NRS	Numerical Rating Scale
NSS-EC	National Statistics Socio-Economic Classification
NZ	New Zealand
OA	Osteoarthritis

OCD	Obsessive-compulsive disorder
OR	Odds Ratio
PHQ	Patient Health Questionnaire
POMS	Profile of Mood States
PRIME-MD	Primary Care Evaluation for Mental Disorders
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROG-RES	PROGnostic Research study
PTSD	Post-traumatic stress disorder
QDIS	Quick Diagnostic Interview Schedule
QOF	Quality and Outcomes Framework
QUORUM	Quality of Reporting of Meta-analyses
RA	Rheumatoid arthritis
RCGP	Royal Collage of General Practice
RCT	Randomised controlled trial
REML	Restricted maximum likelihood
RRR	Relative risk ratio
SCID	Structured Clinical Interview for DSM disorders
SD	Standard Deviation
SE	Standard Error
SF	Short Form Health Survey
SIGN	Scottish Intercollegiate Guidelines NHS
SRM	Standardised Response Mean
STAI	State-Trait Anxiety Inventory
The GAD	Generalised Anxiety Disorder scale
UAE	United Arab Emirates
UK	United Kingdom
U.S.A	United States of America
VIF	Variance inflation factor
WHO	World Health Organization
WHO DAS	World Health Organization Disability Assessment Schedule for DSM-IV
WMH WHO-CIDI	World Mental Health Survey Initiative version of the World Health Organization's Composite International Diagnostic Interview

Structure of the thesis



Publications and presentations arising from this thesis

Peer reviewed publications:

Rzewuska M, Mallen CD, Strauss VY, Belcher J, Peat G. The Course of Comorbid Anxiety Symptoms in Patients Presenting to General Practice with Symptomatic Osteoarthritis: A Latent Class Growth Analysis. Rheumatology 2012; 51: 110. **(Abstract)**

Oral presentations:

Rzewuska M, Mallen CD, Belcher J, Peat G. The prevalence of comorbid depressive disorders and depression symptoms in adults with osteoarthritis and/or joint pain: systematic review and meta-analyses. Society for Academic Primary Care Annual Scientific Meeting, Bristol, England, 6 - 8 July 2011

Rzewuska M, Mallen CD, Belcher J, Peat G. Depression and anxiety coexisting with osteoarthritis in primary care: from recognition to management. Guest Lecture at University of Central Lancashire, England, 28 April 2011

Poster presentations:

Rzewuska M, Mallen CD, Peng VY, Belcher J, Peat G. A latent class growth analysis of anxiety symptoms in a longitudinal cohort of primary care patients with symptomatic osteoarthritis. Division of Health Psychology Annual Conference, Southampton, England, 14 - 15 September 2011

Rzewuska M, Mallen CD, Strauss VY, Belcher J, Peat G. The Course of Comorbid Anxiety Symptoms in Patients Presenting to General Practice with Symptomatic Osteoarthritis: Latent Class Growth Analysis. British Society for Rheumatology. Glasgow, Scotland, 1 - 3 May 2012

See Appendix A on page 317 for full conference abstracts.

“Did I know what it was? It was my pain, but I don’t think I would have called it depression. I think I would have called it my pain.”

(Karp, p.13, 1994)

Chapter one: Thesis outline, key concepts and operational frameworks

1.1 INTRODUCTION

Chapter one introduces key concepts to this thesis. Next, the reasons why it is important to investigate depression and anxiety coexisting with OA are discussed. The existing evidence for associations between OA and depression/ anxiety, underlying reasons for this potential association and the impacts on the individual are briefly considered. The approach taken by the National Institute for Health and Clinical Excellence (NICE) evidence-based clinical guidelines on depression and anxiety coexisting with OA is introduced and gaps in guidance (i.e. related evidence) are identified. These form the rationale for this thesis, with the chapter being concluded with the primary aim and objectives of the PhD.

1.2 KEY CONCEPTS AND DEFINITIONS

1.2.1 Osteoarthritis (OA)

Osteoarthritis is the most common type of arthritis (Sacks et al., 2010). OA can develop in any synovial joint, but the knees, hips and small hand joints are most commonly affected (National Collaborating Centre for Chronic Conditions (NCC-CC), 2008). The development of a satisfactory definition of OA has presented substantial difficulties, with some authors even questioning the appropriateness of pursuing a 'definition of clinical osteoarthritis' (Hurley et al., 2007). It has been described as a degenerative non-inflammatory disease that encompasses "*failed repair of joint damage*" and characterised by a linear progression (Lane et al., p. 479, 2011). There is, however, evidence indicating that inflammation can occur in OA (Ambramson, 2003,

Brooks, 2003 in Dieppe & Lohmander, 2005). The linear and fixed progression of osteoarthritis is not always the case, as demonstrated by evidence that OA is a metabolically active and dynamic disease involving the damage of cartilage, bone thickening and formation of new bone (Sahlström et al., 1997, Felson et al., 2000, Sharif et al., 2004). Subsequently, OA has been described as a multifactorial (Dieppe & Lohmander, 2005, NICE, 2008) and heterogeneous group of disorders (Dieppe & Lohmander, 2005). Given this complexity, it is increasingly recognised that OA requires a multifaceted, stepwise, patient-specific and aetiology-related approach to management (Dieppe & Lohmander, 2005, NICE, 2008, Lane et al., 2011).

Structural features on plain radiography have been traditionally used to determine the presence of osteoarthritis (Felson et al., 1997), although the use of more sophisticated imaging (e.g. magnetic resonance imaging) offers the promise of new case definitions of OA (e.g. Hutton & Vennart, 1995, Tan et al., 2005). The value of a definition of OA based solely on radiographic changes is questionable, given the widely reported discordance between symptoms and radiographs (Hannan et al., 2000). Structural OA changes can be evident on radiographs without patients reporting pain; and joint pain can be present without X-ray evidence of OA changes to joints (McAlindon et al., 1996, Dieppe & Lohmander, 2005, NICE, 2008, Lane et al., 2011). Despite this discordance, there is nevertheless, a consistent association between symptoms and severity of radiographic features (Duncan et al., 2006, Zhang & Jordan, 2010).

Traditionally, the American College of Rheumatology (ACR) clinical classification criteria (Altman et al., 1986, 1991) have been used for case definitions in research, although their validity in population studies has been questioned (Schouten & Valkenburg, 1995, Bierma-Zeinstra et al., 1999, Peat et al., 2006b) and there have been several recent attempts to provide evidence-based recommendations on the clinical diagnosis of

OA (e.g. Zhang et al., 2009, 2010). The Osteoarthritis Research Society International Disease State working group and the United States Food and Drug Administration, suggest that OA can be defined either as a clinical syndrome or a pathological disease (Lane et al., 2011). The clinical signs and symptoms of OA include use-related joint pain, joint stiffness and bony swelling associated with substantial functional limitation (NICE, 2008). In older adults osteoarthritis is the most common reason for joint pain (McCormick et al., 1995), and thus, diagnosing joint pain as osteoarthritis is increasingly valued in research (Peat et al., 2005) and a clinical syndrome definition is used in NICE OA clinical guidelines (NCC-CC, 2008, NICE, 2008).

1.2.2 Depression

The term depression is used to describe “a mood, a symptom and a syndrome” (Mendels, 1970). There is a lack of agreement on definition (Pilgrim & Bentall, 1999). In psychological theories it refers to emotional state (Lazarus, 2006), but for medical purposes has been defined as a syndrome with an underlying mood disorder that requires “the presence of several symptoms” (Montgomery, 1990).

Depression is characterised by feelings of sadness (low/depressed mood) accompanied by hopelessness, loss of interest in previously enjoyable activities (anhedonia), sleep and appetite disturbances, feelings of worthlessness and, in some cases, thoughts of death (Feliciano & Arena, 2007). It has been argued that high levels of negative affect, for example, low mood, and low levels of positive affect, such as anhedonia, are key aspects of depression (Clark & Watson, 1991, Watson et al., 1995, Marshall et al., 2003, Cook et al, 2004). Affective symptoms are considered in combination with somatic and cognitive symptoms (Turk & Okifuji, 1994). Somatic symptoms of depression include: fatigue or loss of energy, significant weight loss/gain, sleep disturbances,

psychomotor agitation and social withdrawal. Cognitive symptoms of depression encompass: negative thinking about oneself or world or others, feelings of worthlessness, diminished ability to think or concentrate or make decisions, and suicidal thoughts (Feliciano & Arena, 2007).

Categorical systems of mental disorder classification are surrounded with criticism related to the ambiguity of indicators of normality and the common comorbidity of disorders (Kendell, 1975, Widiger & Sankis, 2000, Widiger & Samuel, 2005), including often coexisting depression and anxiety (Cameron, 1985). A categorical approach is considered to be psychopathology-focused and is traced to Kraepelin (1917, cited in Pilgrim & Bentall, 1999), who as a medical naturalist, assumed an existence of "*a real and invariant external world of natural disease entities*" (Pilgrim & Bentall, p. 261, 1999). Currently, a categorical approach is used and can be partially attributed to several factors including tradition, credibility, and endeavoured simplicity, utility and validity (for a summary see Widiger & Mullins-Sweatt, 2007).

In 1948, the World Health Organization (WHO) revised the Manual of the International Statistical Classification of Diseases Injuries and Causes of Death (ICD-6) and added a mental disorder section (WHO, 1948). The Diagnostic and Statistical Manual of Mental Disorders, First Edition (DSM-I) was the first diagnostic system specific to mental disorders, developed in 1952 by the Committee on Nomenclature and Statistics of the American Psychiatric Association (APA, 1952). These two classifications remain the most commonly used systems, with the DSM-IV-TR (APA, 2000) and ICD-10 (WHO, 1992) versions currently available. The ICD-10 system of coding is officially used in clinical practice in the UK. Given evidence for diagnostic discordance (López-Ibor et al., 1994, Andrews et al., 1999, Slade

& Andrews, 2001, Andrews & Slade, 2002, López-Ibor, 2002, Vilalta-Franch et al., 2006, First, 2009), caution in comparing estimates, based on those two diagnostic classifications, is required.

In the ICD-10 Classification of Mental and Behavioural Disorders (WHO, 1992), depression is classified as mood disorders including depressive episode, recurrent depressive episode and dysthymia. Mixed anxiety and depressive disorder belongs to a group of neurotic, stress-related and somatoform disorders (WHO, 1992). The DSM-IV-TR also classifies depression as mood disorders including: major depressive disorder, dysthymic disorder, depressive disorder not otherwise specified (APA, 1994). The DSM-IV-TR (APA, 2000) also includes adjustment disorders characterised by the presence of a more severe depressive mood than expected in a reaction to distress. In both classifications the presence and type of disorder is decided based on a number of symptoms, their duration and severity, and clinical significance, which encompasses the presence of distress and disability (Gruenberg et al., 2005). Both the DSM-IV-TR and ICD-10 provide psychotic and remission specifiers, but some course specifiers, such as a seasonal characteristic, are included in the DSM-IV-TR, but not in the ICD-10 (Gruenberg et al., 2005).

1.2.3 Anxiety

Similarly to depression, there are disparities about what constitute the definition of anxiety. Efforts have been made to build a consensus (Whitley, 1992, Bay & Algase, 1999). Anxiety can range from a temporary response to the maladaptive experience of prolonged and an intensive anxiety syndrome with an underlying anxiety disorder (Sarason & Sarason, 2004). The term anxiety has been used in animal models to describe an emotional state of heightened sense of apprehension and hypervigilance in response to an unrecognisable threatening stimulus (Blanchard &

Blanchard, 1989), with an underlying disassociation between the source and response (LeDoux, 1996). It has been hypothesised that information mismatch is translated into emotional, cognitive, behavioural, and somatic response including (Blanchard & Blanchard, 1989, Spielberger & Rickman, 1991, Seligman et al., 2001, Sarason & Sarason, 2004, Beesdo et al., 2009).

- Emotional symptoms (excessive worries and a feeling of uncertainty)
- Cognitive symptoms (sense of confusion, poor concentration, negative thoughts)
- Behavioural symptoms (indecision, avoidance)
- Physiological symptoms (muscular tension, tightness in the neck)
- Hypervigilance symptoms (irritability, restlessness)

Anxiety is believed to share with depression high levels of negative affect, but physiological hyperarousal is considered specific to anxiety (Clark & Watson, 1991). Cardiovascular excitation and apprehension is also a characteristic of fear (Whitley, 1994), as both anxiety and fear involve activation of the sympathetic nervous system (Kalin, 1993, Neumann et al., 2011). Similarities and differences between fear and anxiety have been simultaneously delineated (Bay & Algase, 1999). Biologically, fear is a motivated defensive state that has been shown to originate from the amygdala (LeDoux, 1996, 2003). Consequently, fear has been defined as a temporary 'flight or fight' emotional response to a discrete actual stimulus identified and perceived as threatening (Blanchard & Blanchard, 1989).

Anxiety is a heterogeneous condition, with anxiety symptoms being a primary feature (APA, 2000). The presence and type of disorder is decided based on a number of symptoms, their duration and severity, and the presence of distress and disability (Gruenberg et al., 2005). The ICD-10 classification includes organic anxiety disorders, which are not relevant to this thesis and non-organic anxiety disorders.

The latter are classified as neurotic, stress-related and somatoform disorders [APA, 2000]. Examples include, phobic anxiety disorders; agoraphobia with panic disorder, agoraphobia without panic disorder, social phobia, and specific phobia, obsessive-compulsive disorder (OCD), other anxiety disorders; panic disorder, generalised anxiety disorder (GAD), mixed anxiety and depressive disorder, reaction to severe stress, and adjustment disorders; post-traumatic stress disorder (PTSD) and acute stress disorder (ASD) [WHO, 1992].

The DSM-IV-TR classification of anxiety disorders includes generalised anxiety disorder, panic disorder with agoraphobia, panic disorder without agoraphobia, agoraphobia without history of panic disorder, specific phobia, social phobia, OCD, PTSD, ASD, anxiety disorder due to a general medical condition, anxiety disorder due to... [indicate the general medical condition] and anxiety disorder not otherwise specified [APA, 2000].

1.2.4 Determining the presence of depression/anxiety symptoms and depressive/anxiety disorders

A number of methods exist to determine the presence of anxiety and depression. These methods are considered below.

- Self-report measures are designed to be compatible with diagnostic classifications of depressive/anxiety disorders [Feliciano & Arena, 2007], but are not diagnostic measures *per se*. Instead they allow clinicians/researchers to assess the severity and/or number of symptoms [Nease & Malouin, 2003]. Depression measures (e.g. the Nine-Item Patient Health Questionnaire [Spitzer et al., 2006]) and some self-report anxiety measures (e.g. the Beck Anxiety Inventory [Beck & Steer, 1993]) assess more general levels of symptoms. Other anxiety measures (e.g.

the Generalised Anxiety Disorder Questionnaire (Newman et al., 2002)) are designed to assess symptoms of specific anxiety disorders. They differ in length and, based on time required to complete, can be divided into ultra-short (1 to 4 items: <2 minutes), short (5 to 14 items: 2-5 minutes) and long/standard (15 or more: >5 minutes) (Mitchell & Coyne, 2007). Questionnaires also differ in psychometric and diagnostic properties (Williams et al., 2002). A self-report measure involves a patient indicating the degree to which one has experienced certain symptoms in a given time frame. The result is scored by a clinician/experimenter with the overall score reflecting the severity/numbers of depression/anxiety symptoms (Feliciano & Arena, 2007). The identification of symptoms with self-report measures is followed by a confirmation of the diagnosis with a more comprehensive assessment (Feliciano & Arena, 2007).

- Diagnosis is reached upon attribution of a symptom manifestation to underlying causes perceived by the clinician (Andrews & Peters, 1998). A structured diagnostic interview can be used for this purpose; a method valued for offering a deeper understanding of the mental health problem (Gibson, 1998). To ensure consistency and avoid misclassification, structured and semi-structured interviews are usually conducted (Sheehan et al., 1998).

1.2.5 Definitions of screening, case identification, detection, diagnosis recognition

The following section will define some of the key terms related to identification and assessment of depressive and anxiety symptoms used in this thesis:

- Screening: Screening is "*the presumptive identification of unrecognised disease or defect by the application of tests, examinations or other procedures*

which can be applied rapidly. Screening tests sort out apparently well persons who probably have a disease from those who probably do not. A screening test is not intended to be diagnostic." (Wilson & Junger, p. 11, 1968)

- Targeted case identification: Kessler et al. (2005) have regarded the term targeted case-finding of depression in high-risk groups as more appropriate than screening. This distinction was made as according to Kessler et al. (2005) it appears that depression does not formally meet widely accepted criteria for a disease that is appropriate for screening. NICE (2009a, 2009b) depression guidelines refer to targeted case-finding as case identification (identification of a specific disease), hence this term will be used in the current thesis in the context of both depression and anxiety.
- Detection: Clinical detection of a disease implies detection in a clinical setting among persons presenting to a clinician or using medical services (Porta, p.74, 2008). Detection is not a synonym for diagnosis, therefore confirmation of the suspected diagnosis will require additional assessment (probably a formal diagnosis) (Porta, 2008). In research detection appears to be used by academic general practitioners in the context of what has been defined as clinical detection (e.g. Dowrick & Buchan, 1995, Licht-Strunk et al., 2009a). For consistency within this thesis the term detection will be used in the context of primary care attendees considered to be possibly or definitely depressed (or anxious) by the general practitioner (GP) (Dowrick & Buchan, 1995). Similar to Licht-Strunk et al. (2009a), in the context of medical records review, this definition will be extended to evidence of initiation of treatment.
- Diagnosis: Diagnosis is the process of "*determining health status and the factors responsible for producing it; may be applied to an individual, family,*

group, or community. The term is applied both to the process of determination and to its findings." (Porta, p. 66, 2008)

- Recognition: The term recognition does not appear to be clearly delineated in health literature and seems to be closely related to detection. In research it seems to be used in two contexts; recognition (identification) of symptoms (e.g. Baik et al., 2005) and recognition of a condition (evidenced by a formal diagnosis/definite problem (e.g. NICE, 2009b)), which typically implies meeting formal diagnostic criteria.

1.3 WHY IS IT IMPORTANT TO INVESTIGATE DEPRESSION AND ANXIETY COEXISTING WITH OA IN PRIMARY CARE ADULTS?

1.3.1 Prevalence of OA, depression and anxiety in primary care

Osteoarthritis

Osteoarthritis (OA) is a worldwide public health problem with an increasing prevalence related to the aging population (Felson et al., 2000, Kopec et al., 2008). It is one of the commonest chronic conditions managed in primary care (ARMA, 2004, ARC, 2002) accounting for an estimated 15% of musculoskeletal disease consultations in general practice (NCC-CC, 2008). As shown by the annual 2007 report for Weekly Returns Service of Royal College of General Practice ((RCGP) 2007), consulting prevalence rates for osteoarthritis are high. They ranged from 252 in 10,000 registered adults aged 45-65 years to 923 per 10,000 registered adults aged 75 years or over (RCGP, 2007).

Depression and anxiety in primary care

Depressive and anxiety disorders are known to be the most common mental health diagnoses in primary care settings [(Goldberg & Lecrubier, 1995, Sartorius et al., 1996)]. According to the annual 2006 RCGP (2006) report, depression and anxiety states are common reasons for consultation in age groups also consulting for OA. Consulting prevalence rates of depression ranged from 70 per 10,000 registered males aged 65-74 years to 202 per 10,000 registered females aged 45-64 years [(RCGP, 2006)]. In these age groups, person consulting prevalence rates of anxiety states were even higher, ranging from 111 per 10,000 registered males aged 65 - 74 years to 320 per 10,000 registered females aged 45-64 years [(RCGP, 2006)].

1.3.2 Coexistence of OA, depression and anxiety

Empirical evidence of the coexistence

The prevalence of OA is known to increase with age [(Woolf & Pfleger, 2003, NCC-CC, 2008)] and multimorbidity, including psychopathologies, are common in older adults [(Marengoni et al., 2011)]. Community-based and primary care studies have consistently shown that adults with chronic painful conditions are at an increased risk of developing depressive and anxiety symptoms [(Sartorius et al., 1996, Gureje et al., 1998, Cole & Dendukuri, 2003, Katon et al., 2007, Moussavi et al., 2007)]. Based on a population of 5438 individuals across 15 primary care centres, the WHO has estimated that depressive disorders were six times more likely in patients with two or more chronic physical health problems [(Sartorius et al., 1996)]. The same data indicated that patients with persistent pain might be four times more likely to have anxiety or depressive disorders than those without pain [(Gureje et al., 1998)]. Through a systematic search of the MEDLINE electronic database, Katon et al. (2007) have investigated the association of

depression and anxiety with chronic medical illness. The review found that patients with arthritis have a high rate of affective disorders (20–40%) (Katon et al., 2007). Surprisingly, "[in contrast to rheumatoid arthritis] *less research has been completed describing the link between OA and psychosocial variables*" (Katon et al., p.152, 2007).

Why OA, depression and anxiety coexist?

Mechanisms underlying OA, depression and anxiety seem unclear and difficult to understand - given their relationships are likely to be bi-directional (Arola et al., 2010). Potentially useful information on related factors can be drawn from psychological models of adjustment to pain - based on or tested in populations with chronic pain (mainly or exclusively originating from musculoskeletal diseases). A shared view of these models is that clinically significant anxiety and depression symptoms coexisting with pain, are a result of difficulties with adjusting to pain.

Under the framework of the fear-avoidance model a cognitive interpretation of pain as threatening (pain catastrophising) leads to pain-fear, which affects attention leading to avoidance behaviours, followed by hypervigilance of bodily sensations, disability, disuse, and depression, which then further fuel catastrophic thinking (Vlaeyen et al., 1995). This model has been supported in a large sample of patients with chronic pain, although older patients had lower levels of pain fear than middle-aged adults (Cook et al., 2006). Heuts et al. (2004) supported a significant association between pain severity, pain-fear and the level of daily functioning in OA. In the updated version of the model, an anxiety pathway was added between pain-fear and pain-avoidance (Asmundson et al., 2004). Fear-avoidance and anxiety were

found to be independently associated with poor functioning in people with OA, but depression was found to influence functioning only when anxiety was low (i.e. anxiety was concluded to be the main factor) (Scopaz et al., 2009).

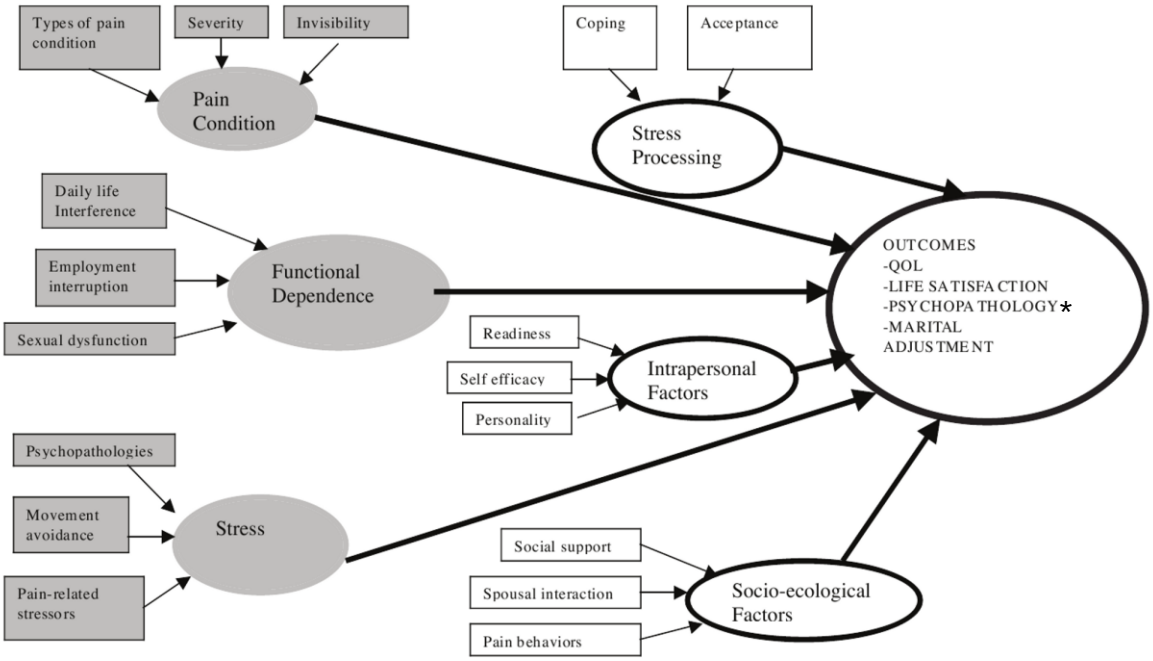
Pain catastrophising plays a key role also in the misdirected-problem solving model (Eccleston & Crombez, 2007). In this model, worries about pain and cognitive evaluation (e.g. pain catastrophising) are adaptive reactions that aim to solve problems, i.e. make sense of pain by searching possible causes, consequences of pain and possible actions. This process engages attention - exaggerating worries about negative consequences and produces hypervigilance to pain. This results in repeated efforts to solve the pain problems (Eccleston & Crombez, 2007). If pain is framed as solely a biomedical problem - it will inevitably lead to attempts to remove or reduce pain. Failed attempts to achieve it further reinforce worries and the person can become stuck in a “perseverance loop” of increasing worries, where the unsuccessful strategy gets repeated. Examination of a small sample of patients with spinal pain showed the link between catastrophising, problem framing, and problem-solving, where the way people viewed their pain problem set the stage for catastrophic worry - a direction more similar to that of the fear-anxiety-avoidance model (Flink et al., 2011). In line with the model, in patients with chronic pain worrying, but not problem solving, had a unique contribution in explaining depressive mood in people with chronic pain (De Vlieger et al., 2006).

In the acceptance-commitment model depression and anxiety are also related to difficulties in adjustment, which stem from dysfunctional cognitive structures (Hayes et al., 1999). In this model anxiety and depression are a result of psychological inflexibility (caused by avoiding or controlling pain) instead of focusing on achievable goals (advocated also in the misdirected problem-solving

model) (Ciarrochi et al., 2010). Pain acceptance was found to moderate the relation between pain and negative affect in a small group of females with osteoarthritis (Kratz et al., 2007). These results were confirmed by McCracken et al.'s (1999) and Vowles and McCracken's (2008) studies with tertiary care patients with chronic pain, where pain-related acceptance led to less anxiety and depression symptoms and higher physical functioning.

All three models acknowledge the impact of predisposing factors. Indeed, in addition to cognitive-behavioural factors, psychological adjustment in people with chronic pain (including musculoskeletal origin) can be affected by a range of external and internal factors (Lee & Mercurio-Riley, 2009). Lee and Mercurio-Riley (2009) reviewed empirical evidence on contributing factors that affect the psychosocial adjustment (anxiety, depression and substance abuse) among individuals with chronic pain (mostly originated in musculoskeletal disease). Following adaptation of the Risk and Resistance Model of Adjustment (Wallander et al., 1989, cited in Lee & Mercurio-Riley, 2009) the authors generated a conceptual framework with various risk and resistance factors, which can be moderated and/or mediated by the resistance factors (see Figure 1.1 overleaf). Variables from each concept were tested by Lee and Chan (2007) for associations with each other and with depression symptoms in people with musculoskeletal pain. Structural pathway modelling showed that life interference, catastrophising, social and family support and stress, directly predicted depression symptoms. Life interferences were found to be predicted by pain severity and coping strategies. Stress was predicted by pre-injury and psychopathology.

Figure 1.1 The conceptual framework of risks (grey denotes) and resistance (white denotes) factors in adjustment to chronic pain.



Source: Lee and Mercurio-Riley, 2009

Note: The rectangle denotes the specific variable which contributes to the concept; * -Typically depression, anxiety or alcohol abuse.

1.3.3 The impact of depression and anxiety on OA

Anxiety and depression symptoms are known to have adverse effects on well-being of patients with OA. Qualitative and quantitative results from a small UK study of primary care patients with OA, found that depressive and anxiety symptoms were listed as one of the most upsetting aspects of OA the experience (Tallon et al., 2000). There is ample evidence from a large community-based study (Arbabzadeh-Bouchez et al., 2002, Tylee et al., 1999) and a large medical outpatients study (Wells et al., 1989b), to suggest that depression symptoms are associated with increased disability in people with arthritis. In a large community-based study depressive episode coexisting with arthritis was found to reduce general health status, more than arthritis alone, depression alone, and any combination of chronic diseases

without depression [Moussavi et al., 2007]. A study of 621 patients with OA found that depression symptoms even increase the risk of developing symptomatic OA in patients with minimal to moderate severity of radiographic changes [Kim et al., 2011]. The impact of anxiety on OA-related outcomes has been less researched, but a study with 182 institutional practice patients with OA found that anxiety symptoms, but not depression symptoms, were associated with diminished physical functioning in OA [Scopaz et al., 2009].

Evidence suggests that both depression and anxiety symptoms may also exert long-term effects on people with OA. A prospective study of 2,558 older primary care patients with chronic physical illnesses (including arthritis) found that patients with elevated depression symptoms at baseline had significantly lower quality adjusted life years over the 4-year study period [Unützer et al., 2000]. In a study of 621 community-based older adults with joint pain, multivariate analyses showed that anxiety symptoms predicted poor functional outcomes at 18 months, but no significant effect was found for depression symptoms [Mallen et al., 2007].

Consequently, a need for an increased attention to depression and anxiety coexisting with OA has been highlighted in numerous studies [Summers et al., 1988, Salaffi et al., 1991, Creamer et al., 2000, Memel et al., 2000, Lin et al., 2003, Mallen et al., 2007, 2008, Axford et al., 2010, Kim et al., 2011]. The importance of recognising and managing depression and anxiety in patients with arthritis has been recently reiterated by the King's Fund and Centre for Mental Health group [Naylor et al., 2012]. A multifaceted approach involving the identification and management of modifiable factors, such as coexistence of depression symptoms and other life stresses have been recognised and incorporated in OA NICE guidelines [NCC-CC, 2008, NICE, 2008].

1.3.4 Evidence-based clinical guidelines for primary care management of depression and anxiety in OA and gaps in guidance

OA guidelines

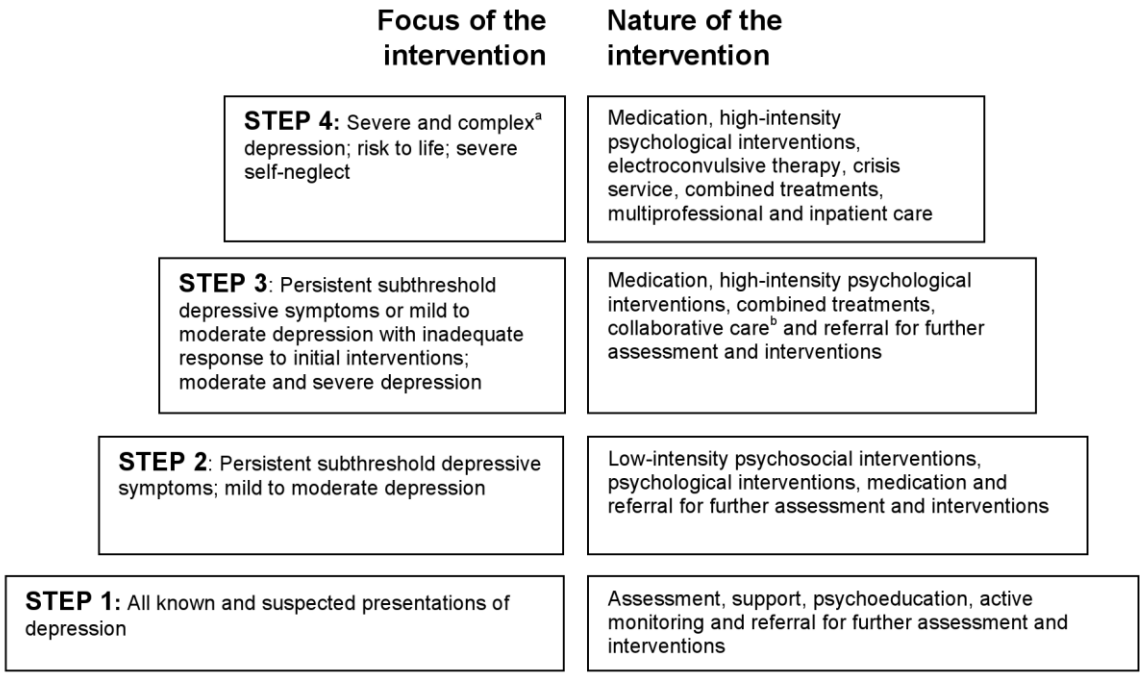
Clinical guidelines are intended to be the first reference source of evidence-informed clinical decision-making in UK primary care (Hunsley & Mash, 2010). The full version of NICE OA guideline cites eleven studies on the association between depression, anxiety and life satisfaction in patients with OA, and recommends a holistic health approach to the management of OA including the active recognition of depression (NCC-CC, 2008). The proposed method of recognition is to "*screen for depression*" in patients with OA (NCC-CC, p. 20, 2008). The guideline provides no details of the best way of screening for depression or the effectiveness of interventions to reduce such depression. Instead it refers to a general guideline for depression (NICE, 2004). Anxiety recognition or management are not mentioned in NICE OA guidelines.

Depression guidelines

Since the publication of NICE OA guidelines there has been a revision to the NICE depression guideline (NICE, 2009a) and the publication of new guidance on the management of depression in patients with physical health problems (NICE, 2009b). While the recommendations in the updated NICE depression guideline are broadly comparable with those made in the previous version, the term case identification (screening in groups at risk) is now used instead of screening. For specific recommendations for patients with chronic physical health problems, the general depression guideline (NICE, 2009a) refers to the depression guideline specific to physical health problems (NICE, 2009b), including diabetes, heart disease, cancer, musculoskeletal disease and respiratory or neurological disorders. Both guidelines

recommend the stepped-care approach to depression care through recognition (awareness, screening and assessment) and management (intervention and monitoring) (Figure 1.2). Similar recommendations are offered by clinical guidelines of the Canadian Task Force on Preventive Health Care (MacMillan et al., 2005) and The US Preventive Services Task Force ((USPSTF), 2002, 2009).

Figure 1.2 The stepped-care model of depression management.



[a] Complex depression includes depression that shows an inadequate response to multiple treatments, is complicated by psychotic symptoms, and/or is associated with significant psychiatric comorbidity or psychosocial factors.
[b] Only for depression where the person also has a chronic physical health problem and associated functional impairment (see 'Depression in adults with a chronic physical health problem: treatment and management' [NICE clinical guideline 91]).

Source: NICE, Depression in adults with a chronic physical health problem: Treatment and management, p.17, 2009b

The recognition of depression involves case identification using two questions, followed by a diagnostic assessment of depression and psychosocial factors, which can be informed using validated measures (NICE 2009a, 2009b). The

Quality and Outcomes Framework (QOF) component of General Medical Services Contract (BMA, GPC 2009) supports NICE's (2009a) recommendations, considers empirical evidence and suggests specific depression measures suitable for using in general practice. It recommends routine depression case identification as part of annual diabetes or coronary heart disease reviews, but not an annual OA review (BMA, GPC 2009). The NICE guideline for depression, coexisting with chronic conditions, advocates asking three additional questions on feeling worthless, poor concentration and suicidal ideations (NICE, 2009b).

Figure 1.2 on page 18 presents a brief summary of recommended interventions based on the focus of depression intervention. A type of intervention should be performed in a framework of the person-centred care (NICE, 2009a). It should also be decided based on patient preferences, the duration and progress of symptoms, previous responses to treatment and possible effectiveness of treatment (NICE, 2009a). According to the guideline for adults with depression and a chronic physical health problem, the collaborative care approach should be considered for patients with moderate or severe depression and coexisting physical health problems (NICE, 2009b). The collaborative care approach encompasses delivery of depression interventions by cooperation between a patient, a GP, a care manager and a psychiatrist to achieve the best approach to management (NICE, 2009b).

Anxiety guidelines

The NICE depression guideline recommends considering treatment of an anxiety disorder first, if the patient has anxiety disorder and coexisting depressive disorders or depression symptoms (NICE, 2009b). The list of relevant anxiety

guidelines includes the Obsessive Compulsive Disorder (OCD) (NICE, 2005), the Post Traumatic Stress disorder guidelines (PTSD) (NCC-MH, 2005) and the Generalised Anxiety Disorder (GAD) and panic disorder (with or without agoraphobia) (NICE, 2007). These three guidelines offer no information on recognition and management of anxiety disorders, relevant to primary care patients with OA.

1.3.5 The importance of improving the understanding of recognition of depressive and anxiety symptoms in primary care patients with OA

There appears to have been less specific advice offered on management of patients with OA and the problem of depression. Currently, advice on OA coexisting with the problem of anxiety seems to be lacking entirely. One cannot confidently assume that arguments and evidence developed in other conditions can be simply applied to OA, as they seem to differ in characteristic, aetiology, progress, impacts on the well-being and available treatments. These factors may affect adjustment processes and thus emotional regulation can be expected to differ across conditions.

NICE depression guidance suggests that management of depression is a process, involving targeted case identification to recognise symptoms and assessment to recognise the problem of depression, with the nature of intervention dependant on the characteristic of depression problem (NICE, 2009b). Awareness of the scope of the additional problems caused by depression and anxiety in patients with OA is important to recognition of depression and anxiety problems, as it can indicate a need for targeted case identification and prevention efforts. Consequently, it helps to inform service provision (Anderson et al., 2001, Barnard et al., 2006). Identification of patients presenting with symptoms of depression or anxiety is

crucial to initiate diagnosis. According to Katon et al. (2007), optimising the management of somatic symptom burden in chronic medical illness is closely related to the accuracy of diagnosis coexisting anxiety and depressive disorders.

A large randomised controlled trial (RCT) of older primary care patients with arthritis found that systematic depression care can reduce not only depression symptoms, but also pain severity, improves functional outcomes and quality of life (Lin et al., 2003, 2006). Similarly, a large RCT in primary care adults found that systematic anxiety care for patients with recognised anxiety disorders can reduce not only symptoms of anxiety, but also depression symptoms and improve functional status (Roy-Byrne et al., 2010). In contrast, an older RCT of primary care patients in the UK found that GPs' awareness of depression symptoms is unlikely to improve prognosis of these patients at 12 months (Dowrick & Buchan, 1995). Primary care health professionals have expressed difficulties with recognition and management of depressive and anxiety disorders in older adults and with chronic physical illness (Burroughs et al., 2006, Murray et al., 2006, Van Rijswijk et al., 2009, Coventry et al., 2011). Previous studies showed that the complexity of recognition and subsequent management of clinically significant depression symptoms in older primary care attendees is reflected in older adults often being undetected (Licht-Strunk et al., 2009a) and untreated (Kendrick et al., 2009) by their GPs.

1.4 SUMMARY

Disagreements exist regarding what constitutes osteoarthritis, but in general practice diagnosis of OA is most often clinical (examination and information on history and other symptoms). Depression and anxiety are descriptive terms, used for two distinct emotional reactions/states, which can

develop into clinically significant symptoms, traditionally classified into heterogeneous disorders. A number and severity of symptoms of depression and anxiety can be recognised through assessment with self-report measures. Recognition of depression and anxiety disorders (a traditional approach) involves a formal diagnosis.

OA is a common reason for consultation in people aged 45 and above, and depression and anxiety disorders are common reasons for consultation in the same age group. It seems well-established that physical chronic illnesses commonly coexist with depression and anxiety. Several modern psychological theories of chronic pain attribute this to difficulties with emotional adjustment to pain. Both depression and anxiety are known to have detrimental consequences for persons with OA. Nevertheless, there appears to be less specific investigations of depression, anxiety coexisting with OA specifically than in other conditions (e.g. cardiovascular disease and diabetes). Furthermore, if evidence on coexisting depression and OA appears to be relatively scarce, evidence for anxiety is even scarcer. Evidence is needed to offer specific clinical guidance on how to recognise primary care patients with OA who may benefit from an intervention for their problems of depression and anxiety. This seems essential because emerging evidence suggests that the potential benefit of managing depression and anxiety problems coexisting with OA is important in itself, but their management could even improve physical functioning, disability and overall quality of life of patients with OA.

1.5 THESIS AIMS AND OBJECTIVES

The overall aim of this thesis is to advance understanding of coexisting anxiety and/or depression in patients with osteoarthritis. The specific aims and objectives addressed to achieve the overall goal of this thesis are summarised below.

1) The **first aim** is to advance understanding of the prevalence of coexisting depressive and anxiety symptoms/disorders in patients with OA.

Specific objectives are:

- To summarise the scientific evidence of prevalence rates of depression/anxiety and depression/anxiety disorders coexisting with OA/joint pain, by systemically reviewing the existing prevalence studies in community-dwelling/primary care adults with OA/joint pain and coexisting depression and anxiety
- To examine sources of between-study variance in reported prevalence rates

2) The **second aim** is to understand the comparative strengths and weaknesses of recommended patient-reported, self-complete, condition-specific measures used in assessing anxiety and depression symptoms in patients presenting to primary care with osteoarthritis.

Specific objective is:

- To conduct a narrative synthesis of evidence on key measurement properties and characteristics and suitability of selected depression (BDI (BDI-II and BDI-

PC versions), HADS-D, PHQ (2- and 9-item versions)) and anxiety (GAD (2- and 7-item versions), HADS-A) symptom measures

3) The **third aim** is to advance understanding of persistence of depressive and anxiety symptoms in older primary care patients with OA.

Specific objectives are:

- To describe changes in the rate of HADS defined depressive and/or anxiety symptoms in older patients presenting to general practice with musculoskeletal pain
- To identify discrete 12-month post-consultation trajectories of symptoms of anxiety and depression in older patients presenting to general practice with musculoskeletal pain
- To explore patterns of coexisting trajectories of symptoms of anxiety and depression symptoms
- To examine their unique relationships with baseline person-related characteristics

4) The **fourth aim** of this thesis is to improve the understanding of the detection of depression and anxiety in older primary care patients with OA.

Specific objectives are:

- To estimate the detection rate of persistent depression or anxiety symptoms in older patients presenting to general practice with musculoskeletal pain
- To establish factors associated with detection in this sub-population

Chapter two: Prevalence of depression and anxiety in adults with osteoarthritis/joint pain: a systematic review and meta-analyses

2.1 INTRODUCTION

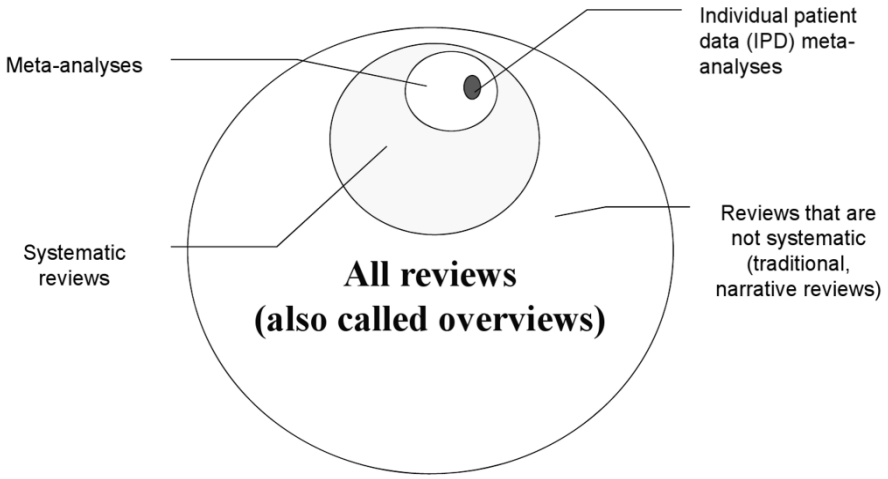
The potential importance of coexisting anxiety and depression to the effective primary care management of osteoarthritis was introduced in chapter one. Chapter two is a systematic review of observational studies investigating the prevalence of depressive and anxiety disorders and questionnaire-assessed depression and anxiety symptoms in primary care patients and community-dwelling adults with osteoarthritis. The chapter starts with introducing the method used and arguing why such a review is needed, before describing the primary objectives. Finally, the results of the synthesised evidence are presented and discussed, and conclusions drawn.

2.2 BACKGROUND

2.2.1 Systematic review and meta-analyses

Literature reviews typically involve narrative summaries of evidence conducted in an informal and subjective manner (Pai et al., 2004). By contrast, a systematic review is "*a review in which there is a comprehensive search for all relevant studies on a specific topic, and those identified are then appraised and synthesised according to a predetermined and explicit method*" (Klassen et al., p. 700, 1998). As depicted in Figure 2.1 overleaf, a systematic review can be followed by a meta-analysis; "*defined as the statistical combination of at least two studies to produce a single estimate of the effect*" (Klassen et al., p. 700, 1998). A meta-analysis can be considered accurate only when resulting from a systematic review (Egger et al., 2001a).

Figure 2.1 Types of reviews.



Source: Pai et al., 2004

A systematic approach is recommended to manage the diversity of available scientific evidence [(Khan et al., 2001, Glasziou, 2001)]. The major advantage of this method is that by stating clear objectives, pre-determined explicit eligibility criteria and a search strategy, allows a more effective and objective identification, appraisal and summary of study findings in the subject of interest [(Egger et al., 2001a)]. Overall, it has been argued that a systematic review reduces the possibility of information and selection bias, making conclusions more objective, accurate and reliable [(Glasziou, 2001)]. In contrast, non-systematic review methods are regarded more predisposed to bias and errors [(Egger et al., 2001a)]. Nevertheless, systematic reviews are time consuming and have the potential for bias if not rigorously designed and executed [(Glasziou, 2001, Pai et al., 2004)], the risk of which is believed to be minimised by following a general guideline such as the PRISMA statement (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [(Liberati et al., 2009)]. See Table 2.1 overleaf for a summary of advantages and limitations of a systematic review.

Table 2.1 Advantages and disadvantages of a systematic review.

Advantages:
• clearly described objectives and eligibility criteria • a balanced picture ensured by inclusion of a broad range of studies • a thorough search • reduced possibility of selection bias (by a verifiable and replicable search strategy and multiple reviewers) • concise and readable format • assessment of methodological quality- indicative of sources of bias in individual studies
Disadvantages:
• oversimplification of differences between individual studies • required considerably more effort than traditional literature reviews • misidentification of relevant literature including unpublished citations • can be misleading by ignoring issues of quality in interpretation

Note: Based on Glasziou, 2001, Pai et al., 2004.

A meta-analysis can be a useful addition to a systematic review, as an alternative to highly subjective narrative synthesis (Egger et al., 1998, Egger et al., 2001a, Pai et al., 2004). Meta-analysis of observational studies (including cross-sectional, longitudinal and case-controls) has become increasingly common in the past 4 decades (Stroup et al., 1997, cited in Stroup et al., 2000). It can involve combining individual measures of the relationship between the frequency of the event in the intervention/exposure group and that in the control/reference group, where the parameter of interest is an odds ratio or risk/rate ratio (Deeks et al., 2001). Alternative to an occurrence relationship, a meta-analysis can focus on occurrence estimates, where the parameter of interest is a prevalence or incidence (Egger et al., 2001a). This type of meta-analysis does not involve calculating effect size, but aggregating an estimate weighted by its precision (Egger et al., 2001a).

Amongst the general advantages of combining and weighting estimates

from several studies are: (a) enabling recognition of inconsistencies in reported estimates (Blettner et al., 1999), (b) generating more accurate estimates than those obtained from any one primary study (Rosenthal & DiMatteo, 2001) and (c) improving comparability to previous research (Rosenthal & DiMatteo, 2001). Nonetheless, a meta-analysis of findings from observational studies, with different aims and methods, may be at a high risk of biased results (Shapiro, 1994, Lipsey & Wilson, 2001). As a result, ways of improving the reliability of meta-analysis of observational studies have been proposed, including: generating a study protocol, conducting a systematic review, acknowledging study quality and biases, and ensuring study replicability (Egger et al., 1998, Egger et al., 2001a, 2001b). Shapiro (1994) has highlighted that high heterogeneity of estimates are likely in meta-analyses of observational studies, which can undermine the accuracy of this method. High heterogeneity of prevalence rates of depression and anxiety have been consistently found across previous meta-analyses (e.g. Anderson et al., 2001, Rutledge et al., 2006, Luppa et al., 2010, Mitchell et al., 2011). In response to this problem, Egger et al. (1998) have suggested that rather than elimination of heterogeneity in meta-analyses of observational studies, a careful examination of its sources is necessary.

Overall, the task of managing meta-analyses of prevalence rates derived from observational studies seems challenging. General methodological issues raised in previous reviews include managing: non-normal distribution of data, heterogeneity and study quality, methodological differences in study design, incompleteness of data and missing information (Polanczyk et al., 2007, Uthman, 2008, Fazel et al., 2008, Luppa et al., 2010, Mitchell et al., 2011). Examples of previously reported challenges specific to meta-analyses of prevalence rates of depression and anxiety include variability in depression and anxiety ascertainment methods and ways of defining

depression and anxiety cases (Grigsby et al., 2002, Gilchrist & Gunn, 2007, Mitchell et al., 2011).

Explicit guidance for decision making around meta-analyses of observational studies of prevalence rates is currently lacking. General quality criteria for meta-analyses are available including: QUORUM (Quality of Reporting of Meta-analyses) (Clarke, 2000), MOOSE (Meta-analysis Of Observational Studies in Epidemiology) (Stroup et al., 2000) and PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Liberati et al., 2009).

2.3 RATIONALE OF THE STUDY

This thesis contends that coexisting anxiety and depression are important to the effective primary care management of osteoarthritis, yet despite the availability of suitable assessment tools, depression and anxiety are often undetected in primary care (Kessler et al., 2002, Licht-Strunk et al., 2009a). Furthermore, in one study with 336 UK primary care patients, 60% and 70% patients with depressive and anxiety disorders respectively had unmet need for care (Boardman et al.,2004). Lord Layard, a health economist, stated that 2.75 million people in England visiting general practices are eligible for psychological therapies, of which only 8% receive such therapy (Centre for Economic Performance's Mental Health Policy Group, 2006).

One of the first steps in epidemiological exploration is to examine the scope of the problem of interest and its distribution across sub-populations in which it occurs. This involves consideration of case definition, populations of interest, and obtaining accurate estimates of occurrence (Webb & Bain, 2011). In the context of anxiety and/or depression coexisting with osteoarthritis, such exploration may serve several purposes.

This should help establish whether this is a common problem of general

importance or a relatively rare occurrence [Glass, 2000]. Such evidence may raise awareness of the distribution of the problem of interest within the population of interest. This may help formulate hypotheses about its causes as well as identifying groups of individuals in whom most attention is required. Increased awareness may also be necessary for formulating health policy, particularly prevention efforts. A 'sufficient prevalence' of the disease in a target group is important for considering routine case identification [Grimes & Schultz, 2002, UK National Screening Committee Recommendations, 2009], which is one of the potential approaches to improving the timely recognition of depressive and anxiety disorders in people with OA.

Prevalence rates of coexisting depression and anxiety were systematically searched and synthesised in other long-term conditions, such as diabetes [Anderson et al, 2001, Grigsby et al., 2002, Ali et al., 2006, Barnard et al., 2006] or cardiovascular disease [Rutledge et al., 2006, Yohannes et al., 2010]. In these conditions, recognition of coexisting depression and anxiety is advocated by clinical guidelines [BMA, GPC, 2009, Scottish Intercollegiate Guidelines NHS (SIGN), 2010, NICE, 2010]. In the field of osteoarthritis there are several studies that have shown that both depressive and anxiety symptoms are common amongst primary care patients presenting to their general practitioners with osteoarthritis [Memel et al., 2000, Rosemann et al., 2007, Mallen & Peat, 2008]. Yet, in the absence of agreed case definitions as well as differences in sample frames and study populations, prevalence estimates may differ widely from one study to another [Jacobsen et al., 2006]. This has been consistently demonstrated by systematic reviews in other chronic conditions [Anderson et al, 2001, Grigsby et al., 2002, Rutledge et al., 2006, Yohannes et al., 2010, Mitchell et al., 2011]. A systematic approach to identifying and summarising published prevalence rates of depression and anxiety coexisting with OA is currently lacking.

This chapter presents the results of a systematic review and meta-analyses conducted to ascertain the prevalence of comorbid depressive and anxiety disorders and questionnaire assessed anxiety and depression symptoms in community based adults with OA/ joint pain. Subgroup meta-analyses to explore prevalence variability across study characteristics are reported (i.e. definition of the condition of interest, geographical location and study setting sample size, mean age, gender distribution, method of ascertainment, study quality). Meta-regression analyses to quantify the effect of these characteristics are reported.

2.4 AIM AND OBJECTIVES

The **aim** is to advance understanding of the prevalence of coexisting depressive and anxiety symptoms/disorders in patients with OA.

Specific objectives are:

- To summarise the scientific evidence of prevalence rates of depression/anxiety and depression/anxiety disorders coexisting with OA/ joint pain, by systemically reviewing the existing prevalence studies in community-dwelling/primary care adults with OA/ joint pain and coexisting depression and anxiety
- To examine sources between-study variance in reported prevalence rates

2.5 METHOD

The PRISMA guideline (Liberati et al., 2009) for systematic reviews with meta-analyses was followed and a systematic review protocol was formulated.

2.5.1 Search strategy

Databases

Published literature was searched from inception to the end of August 2009 using the following electronic databases: EMBASE, MEDLINE, PsycInfo, CINAHL, Web of Science and CSA illumina. Unpublished literature was searched from inception to the end of August 2009 using CSA illumina and Web of Science. References cited in identified articles were examined and citations were tracked.

Search terms

A broad search strategy using both text words and thesaurus terms was designed with a help of an information specialist. Key areas searched for included 'depression', 'anxiety', 'primary care', 'general population', 'joint pain', 'osteoarthritis', 'prevalence', 'observational studies'. For the detailed search strategies please refer to Box B.1.1 (in Appendix B.1 on page 323).

2.5.2 Eligibility Criteria

The following inclusion criteria were used in the review:

- Observational study design
- Adult participants
- Primary care or general population settings
- Diagnosed osteoarthritis or symptoms of OA/ joint pain

- Multiple articles on the same cohort were treated as a single study

Excluded from the review were:

- Non-English language citations
- Secondary care populations
- Studies without primary data
- Samples comprising exclusively of patients with inflammatory arthritis (e.g. rheumatoid arthritis), spondyloarthropathy (e.g. ankylosing spondylitis), crystal arthropathy (e.g. gout), or other non-OA musculoskeletal disorders (e.g. fibromyalgia)

2.5.3 Study selection

Titles and abstracts were examined by the first reviewer. The process of initial screening of full texts for eligibility and selection for further consideration involved consensus between two reviewers. Following this, data extracted from selected articles by the first investigator was checked and quality assessed by another reviewer. At this point further exclusions were made on the basis of: duplicate papers, studies for which required data could not be obtained from authors, data for mixed depression and anxiety, and studies in which definitions of OA did not clearly meet the eligibility criteria. Studies with reported mean scores for depression and anxiety, but no prevalence rates were excluded from the systematic review but retained for use in chapter three. All disagreements in the process of study selection were reconciled by consensus with an independent third reviewer.

2.5.4 Data extraction

The following descriptive characteristics were extracted from each of the eligible articles: authors, year of publication, location of study (country, city), setting, study design, response rate, number of participants, assessment method of the musculoskeletal disease of interest, anatomical site, and additional comments (e.g. to record information on variables such as ethnicity of participants and key study findings). In addition, the following information was extracted: depression and anxiety definition and prevalence, psychological assessment tool, mean scores of depression/anxiety symptoms and measure of distribution of scores (depending on availability of data within the identified studies). A data extraction sheet was developed, including columns for the above variables and a column for additional comments on key findings utilised in the discussion. One reviewer extracted data from all included articles and the second reviewer checked the extracted data. Authors were contacted to obtain missing data or to clarify identified studies. Disagreements were resolved by consensus.

2.5.5 Quality assessment

Methodological quality was assessed to identify bias in individual studies. To ascertain the validity of eligible studies, six validated checklists (Loney & Stratford, 1999, Dionne et al., 2006, Mallen et al., 2006a, Pincus et al., 2006, Van Dijk et al., 2006, Thombs et al., 2007) previously used in systematic reviews of musculoskeletal studies formed the foundations of a quality assessment checklist. Additional articles were sought if methodological details were provided in linked publications (e.g. protocol papers). Quality assessment criteria were rated using the 'yes/no/unclear' (or not applicable) method, according to the presence/absence/lack of clarity (Jüni et al., 1999)

(or applicability of each criterion).

For the purpose of inclusion in meta-regression analyses each 'yes' received a score of one, 'no' or 'unclear' were assigned zero points allowing summary quality scores to be generated for each article. Quality ratings were assigned, in relation to the number of criteria that were met, using the following scale: poor (0-7), fair (8-10) and good (11-15). The cut-off points were devised based on previously used systems (Pincus et al., 2002, 2006, Cesario et al., 2006), and can be found displayed below in Table 2.2.

Table 2.2 Quality rating system.

Percentage of criteria met	Quality score	Quality rating
75%-100%	11-15	Good
50%-74%	8-10	Fair
≥ 49%	0-7	Poor

Two reviewers assessed the methodological quality of each study using the quality checklist. The results were discussed at a series of meetings between reviewers. Levels of agreement between the first reviewer and four other reviewers were estimated (Sheskin, 2007) and interpreted (Altman, 1991) using unweighted Cohen's Kappa.

2.5.6 Data presentation

The result section includes: presentation of the selection process, methodological quality assessment, description of studies and narrative synthesis of prevalence rates for which meta-analyses were unfeasible. Following this, results of the meta-analyses are presented including questionnaire data for clinically important levels of anxiety and depression symptoms and clinical

interview data for anxiety and depressive disorders. The result of meta-analyses section concludes with a summary of results. Supplementary analyses are then discussed, including sensitivity analyses conducted to assess impact of individual studies on pooled prevalence rates.

2.5.7 Data analyses

Data entry was completed using SPSS software (version 15). Meta-analyses were conducted with STATA software (version 11.1). Analyses were conducted independently for anxiety and depressive disorders assessed with clinical interviews and anxiety and depression symptoms assessed with self-report questionnaires. Prevalence rates were meta-analysed and the general principles outlined by Sterne et al. (2003) were applied. A meta-analysis (Klassen et al., 1998) was conducted when at least two estimates were available for a specific type of depressive/anxiety disorder. Data not included in meta-analyses was reported in narrative synthesis.

2.5.8 Classification of prevalence estimates

All depressive and anxiety disorders defined by the DSM-IV-TR (APA, 2000) and the ICD-10 (WHO, 1992) diagnostic criteria were included, reported separately and analysed. Meta-analyses of self-report questionnaires were performed for all "clinically relevant" (qualify for an intervention) symptoms of depression and anxiety (typically mild to severe) (Anderson et al., 2001). Literature for specific questionnaires was searched to identify recommended cut-off scores. Self-report questionnaire data was divided into two categories: 1) possible depression/anxiety (mild symptoms or the minimum threshold of clinical relevance established for the

specific tool) 2) probable depression/anxiety (warranting treatment [NICE, 2009b]). The first group was referred to as 'mild or worse' symptoms and the second group was referred to as 'moderate or worse' symptoms. In cases of longitudinal studies and secondary data analyses, reported cross-sectional prevalence rates from the earliest wave of data collection were included.

2.5.9 Overall and sub-grouped meta-analyses

Given the anticipated large degree of heterogeneity, estimates were pooled and explored using a random effects approach [Eagger et al., 2001a]. Normal distribution was required for pooling data, thus logit transformation was applied as outlined and adopted for pooling prevalence estimates by Uthman (2008) for effect size and standard error and weighted by inverse variance of logit transformed prevalence. Once generated, pooled logit estimates were back transformed to proportions [Uthman, 2008].

The pooled estimates were calculated for each group (command *metan*), the precision (95% CI) was determined and heterogeneity assessed, using Cochran's Q test (reported with chi^2-value and p-value) and the I^2 test. The latter provides a measure of the degree of heterogeneity, with low, moderate, and high values of 25%, 50%, and 75% [Higgins et al., 2003].

Sub-group estimates were calculated (command *metan, by (covariate)*) for disorders and 'mild or worse' symptoms. Subgroup meta-analyses involved recommended [Stroup et al., 2000] or previously investigated [Fazel et al., 2008] study characteristics. Pre-defined factors across which pooled estimates and individual studies were compared (depending on the availability of information) included: method of defining OA/ joint pain cases, anxiety/depression assessment tool,

study setting, proportion of females (dichotomised by median), study quality (poor/fair vs. good), age (dichotomised by median), study geographical location and sample size (>200 vs. ≤200 (Fazel et al., 2008)). As prevalence rates for 'moderate or worse' symptoms severity were derived from samples for which 'mild or worse' symptoms were reported, it was regarded as sufficient to perform sub-group analyses for the latter only. Studies with no variability in covariates (e.g. studies of women only) were omitted from the relevant sub-group analysis as were those with more than 50% missing covariate data (Polanczyk et al., 2007). Between-group heterogeneity cannot be calculated if high within-group variance is present (Sterne et al., 2003).

2.5.10 Meta-regression analyses

To quantify the effects of between-study sources of heterogeneity a meta-regression analysis was applied to prevalence estimates as outlined and operationalised by Harbord and Higgins (2008). Random-effect meta-regression analyses were performed (command *metareg*) with Knapp-Hartung modification for coefficients (including the calculation of standard errors, p-values, and confidence intervals). Random-effects meta-regression estimates the between-study variance followed by estimation of the coefficients (using weighted least squares) (Harbord & Higgins, 2008). Restricted maximum likelihood (REML) regression estimate is the default algorithm in *metareg* (Harbord & Higgins, 2008). All covariates included in sub-group analyses were considered. Studies with insufficient variability in covariates (≥70% sample with same value) were excluded. Initially, covariates were tested individually for their associations with prevalence rates (an unadjusted model). All covariates with p≤0.10 in unadjusted models were included

in adjusted (multivariable) models [Fazel et al., 2008], for which statistical significance was established at the level of $p<0.05$.

2.5.11 Test for publication bias

The presence of publication bias was investigated using Egger's test (command *metabias*) as outlined by Harbord et al. (2009). Egger's test is for small-study effects, and involves regressing the standard normal deviate of intervention effect estimate against its standard error [Harbord et al., 2009]. Notably, whilst it is a test for publication bias, in practice the test indicates whether precision in estimating a pooled prevalence increases as the sample size of component studies increases. Therefore, results of this test should not be interpreted as indicative of bias in likelihood of being published *per se*. As such the test was considered here suitable for prevalence rates reported in the same publication (e.g. a multinational study), as long as estimates are derived from different samples.

2.5.12 Sensitivity analyses

To explore impact of individual estimates on the pooled effect, sensitivity analyses were conducted. To estimate the impact of inclusion of prevalence rates derived from samples with 'arthritis', they were omitted and pooled prevalence rates re-estimated.

In addition to between study variance, heterogeneity can also be related to inclusion of specific, biased studies. Exploration of this problem can involve a non-standardised method of investigating the influence of each individual study on the overall pooled estimate and 95% CI [Sterne et al., 2003]. For this purpose *metaninf* command in STATA can be used, with the production of a numerical table (*print*

id), showing the pooled estimate omitting each study in turn. Graphs can also be generated to visually provide the same results in a plot. The non-standardised method of investigating impact of individual studies involves assessment of a relative difference between the omitted meta-analytic estimate and the combined analysis. Recently introduced methods of sensitivity analyses include equally effective, sequential and combinatorial analyses (Patsopoulous et al., 2008). A sequential algorithm involves omitting one estimate at each time (n-1). "*The study that is responsible for the largest decrease in I^2 is dropped and a new set of n-1 studies is created*" (Patsopoulous et al., p. 1149, 2008) combinatorial algorithm involves omitting two studies at a time. The method has been criticised for being potentially misleading (Higgins, 2008). To explore the impact of individual studies, both the commonly used method and the sequential algorithm method, were implemented and compared. Numbers of estimates required to achieve ≤75% and ≤50% I^2 using both methods are reported.

2.6 RESULTS

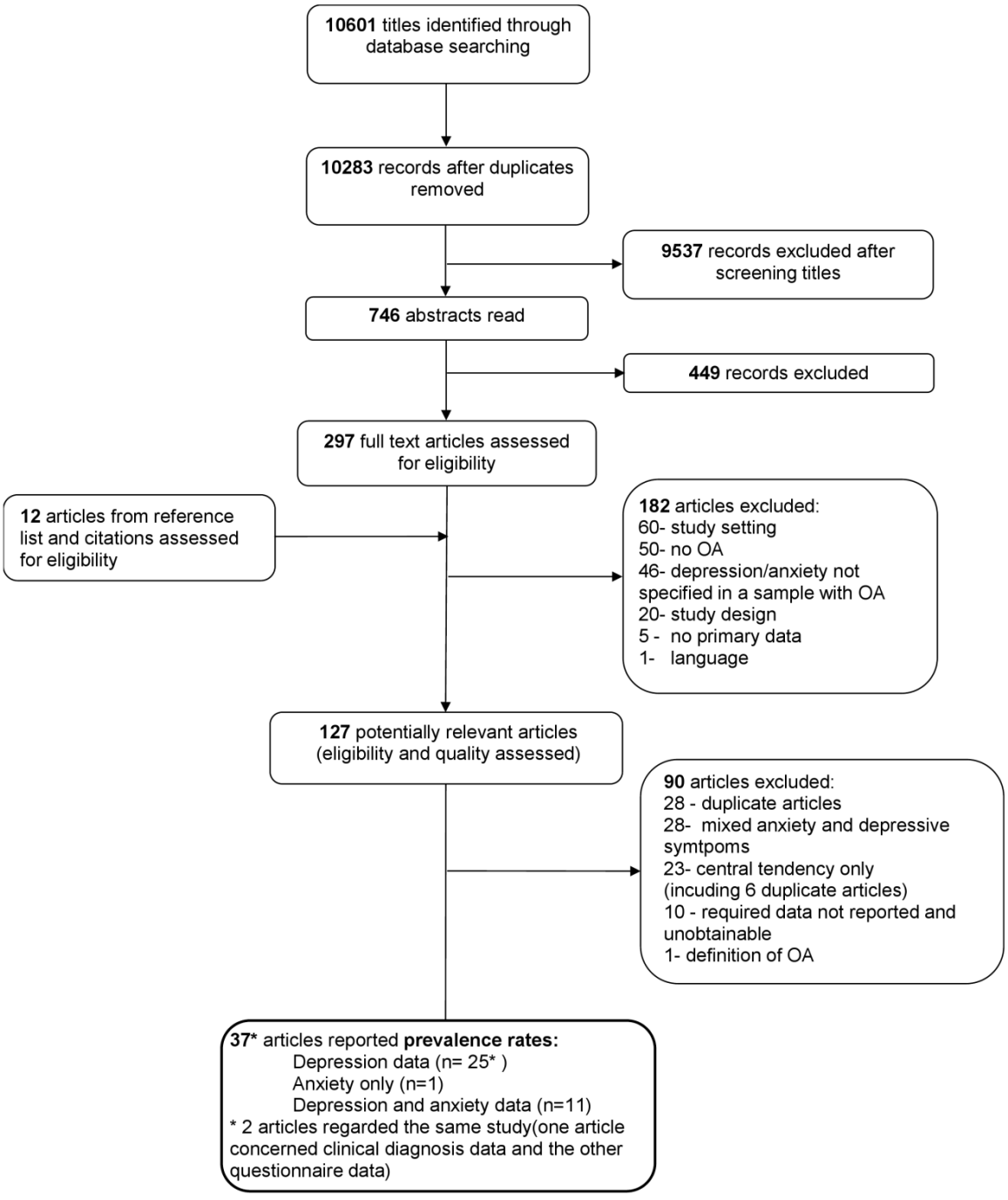
2.6.1 Selection of studies

Table 2.3 Numbers of papers identified at each stage of the search.

	Hits	After removing duplicates	Abstracts read	Initial screening of full text	Text Further screened for eligibility	Included in the review
PsycInfo	361	343	37	17	5	0
CINAHL	713	690	43	16	4	1
Medline	1991	1980	180	84	40	9
EMBASE	3058	2938	82	36	15	4
CSA Illumina	1132	1109	94	48	21	11
Web of Science	3346	3223	310	96	34	11
Total	10601	10283	746	297 (+12 from citations)	119 (+8 from citations)	36 (+1 from citations)

Table 2.3 gives details of papers identified through 6 databases, at each stage of the search process. The literature search identified 10601 articles (10283 after removing duplicates), 9537 papers were excluded on title review and the remaining 746 abstracts were examined. Of these, 309 full text articles were assessed for eligibility and 127 of these were chosen for further eligibility and quality assessment. Included in the current review were 37 articles (36 studies) for which prevalence rates of depression and anxiety were reported or provided by authors upon request (2 cases). One study reported rates for self-report questionnaires and clinical interviews (Dunlop et al., 2004, 2005) and thus both articles were considered to be the same study. For flow of information through stages of the review see Figure 2.2 overleaf and for an overview of the included articles refer to Table B.2.1 (in Appendix B.2 on page 328).

Figure 2.2 Flow of information though the different phases of a systematic search.



2.6.2 Quality assessment

Quality characteristics and total quality scores of each of all 127 articles that were assessed and underwent further screening for eligibility, are provided in Table B.3.1 (Appendix B.3, p. 338). An analysis of agreement on quality score (ranging from 'poor' to 'very good') between reviewers, supports the effectiveness of a checklist and clarity of its criteria are listed in Table 2.4. As shown levels of agreement on quality ratings of the 37 included studies were comparable to those yielded for all 127 potentially relevant studies.

Table 2.4 Inter-rater agreement for all 127 full texts articles and the included 37 articles.

Reviewers initials	Unweighted Cohen's Kappa coefficient	95% CI	SE	Level of agreement on quality rating*
Anxiety and/or depressive data (all 127 full texts articles)				
BN/MR	0.49	0.42, 0.55	0.03	Moderate
GMP/MR	0.63	0.55, 0.72	0.04	Good
CM/MR	0.68	0.60, 0.75	0.04	Good
KB/MR	0.89	0.85, 0.94	0.02	Very good
Anxiety and/or depressive data (37 included articles)				
BN/MR	0.55	0.40, 0.69	0.07	Moderate
GMP/MR	0.67	0.55, 0.79	0.06	Good
CM/MR	0.70	0.59, 0.77	0.05	Good
KB/MR	0.88	0.82, 0.94	0.03	Good

Note: * - Interpretation: <0.20 - poor; 0.21-0.40 fair, 0.41-0.60 moderate, 0.61-0.80 good, 0.81-1.00 very good (Altman, 1991); BN - Barbara Nicholl, CM - Christian Mallen, GMP - George Peat, KB - Kay Benyon, MR - Magdalena Rzewuska.

A summary of issues around specific quality domains are presented for the included 37 articles. The quality of studies described in 2 articles was classed as 'poor', 15 'fair' and 20 'good'. Quality domains that were most frequently addressed in original articles were: the quality of data, analysis, results and data presentation (Table 2.5 overleaf). The quality assessment criteria that were less frequently reported in studies reporting anxiety/depression separately, included:

inclusion of a sample size calculation (reported in n=6); the description of non-respondents (n=12); the description of eligibility criteria (n=13); and clear definition of OA (n=17). These should be regarded as possible sources of bias.

Table 2.5 Quality of studies regarding depression/anxiety described in 37 articles.

Item	Quality Criteria	Exemplar	No. of articles*		
			+	-	?
	Study design:				
A	Is study objective clearly defined?	Mallen et al. (2006a)	35	2	0
B	Is the study design appropriate for the research question?	Loney & Stratford(1999) Van Dijk et al. (2006a)	34	1	2
C	Is the sample size calculation presented?	Loney & Stratford(1999) Van Dijk et al. (2006) Thombs et al. (2007)	6	31	0
D	Are the inclusion/exclusion criteria adequately described?	Mallen et al.(2006)	13	19	5
E	Is definition of the physiological condition of interest (case) clear?	Pincus et al. (2006) Van Dijk et al. (2006)	17	16	4
	Study population:				
F	Are the study subjects obtained appropriately for a given study design?	Loney & Stratford(1999) Pincus et al. (2006) Thombs et al. (2007)	30	2	5
G	Is the setting described in detail?	Loney & Stratford (1999)	26	9	2
	Non-respondents:				
H	Is response rate $\geq$ 70%*****	Pincus et al. (2006) Thombs et al. (2007)	23	9	4
I	Is information about non-respondents provided?*****	Van Dijk et al. (2006)	12	22	2
	Quality of data:				
J	Were all data collected directly from the subjects?	Dionne et al. (2006)	35	2	0
K	Is standardised collection of data used?	Mallen et al. (2006a)	30	2	5
L	Are objective and suitable criteria used for measurement of outcome?	Loney & Stratford (1999) Thombs et al. (2007) Pincus et al. (2006) Dionne et al. (2006)	34	3	0
	Analysis, results and data presentation:				
M	Are frequencies of the primary outcome measures given in detail?	Loney & Stratford (1999) Van Dijk et al. (2006)	35	2	0
N	Are employed analysis techniques appropriate?	Van Dijk et al. (2006)	33	0	4
O	Is description of study participants adequate?	Thombs et al. (2007)	27	10	0

Note: * - one case of N/A; + - yes; - no; ? - unclear.

2.6.3 Included studies

Description of studies

An overview of included studies can be found in Table B.2.1 (in Appendix B.2 on page 328). Thirty six studies (described in 37 articles) reported prevalence estimates of depressive and/or anxiety disorders and/or symptoms. 23 studies used cross-sectional design, 11 (described in 12 articles) prospective-cohort and 2 nested case-control design. Nine studies were conducted in primary care and 27 studies (described in 28 articles) in the general population. Sample sizes ranged from 81 (Nour et al., 2005) to 23405 (Fuller-Thompson et al., 2009). The mean age of participants with OA/joint pain ranged from 35.2 (He et al., 2008) to 83.8 (Jakobsson & Hallberg, 2006) years. The percentage of female participants ranged from 31.5% (Creamer et al., 1999b) to 100% (Szoeke et al., 2008). Studies were conducted in the U.S.A (n=12, described in 13 articles), France (n=1), the UK (n=8), Sweden (n=1), Netherlands (n=2), Canada (n=5), China (n=2), Germany (n=1), Australia (n=1), Nigeria (n=1), a multinational study (n=2).

Definitions of OA could be grouped into the following categories: symptomatic OA (n=9), self-reported OA (n=4), medical records defined OA (n=2), the American College of Rheumatology (ACR) criteria defined OA (n=3), radiographic OA (n=1). Through agreement between reviewers involving careful consideration of characteristics of samples (e.g. age and comorbidities) included are also studies concerning ‘arthritis’ (n=7, described in 8 articles). Inclusion of studies involving samples with arthritis is based on the evidence suggesting that osteoarthritis is by far the most common type of arthritis (Sacks et al., 2010). There is also evidence that the term ‘arthritis’ is being used exchangeable to the term ‘osteoarthritis’ by primary care patients (Peat et al., 2005). These studies will be referred

throughout as ‘arthritis’ and the difference in prevalence estimates between clearly defined OA and ‘arthritis’ was explored in a sensitivity analysis.

Identified prevalence rates

In total 24 eligible studies (described in 25 articles) provided prevalence estimates of depressive disorders and/or symptoms (Barberger-Gateau et al., 1992, Dexter & Brandt, 1994, Woo et al., 1994, Creamer et al., 1999, Scudds & Robertson, 2000, Wilcox et al., 2000, Kramer et al., 2002, Dunlop et al., 2004, Fisher, 2004, Figaro et al., 2005, Dunlop et al., 2005, Nour et al., 2005, Polsky et al., 2005, Jakobsson & Hallberg, 2006, Kadam & Croft, 2007, Moussavi et al., 2007, Muus, 2007, Niti et al., 2007, Rosemann, 2007, Allen et al., 2008, Gureje et al., 2008, Sale, 2008, Schram et al., 2008, Szoeke et al., 2008, Fuller-Thomson & Shaked, 2009). In addition, 11 eligible studies provided prevalence estimates of both depressive and anxiety disorders and/or symptoms (O'Reilly et al., 1998, Memel et al., 2000, Croft et al, 2005, Patten et al., 2006, Peat et al., 2006a, Hill, 2007, Leveille et al., 2007, Wilkie et al., 2007, He et al., 2008, Mallen & Peat, 2008, McWilliams et al., 2008) and one of anxiety disorders only (Wells et al., 1989a). Studies which had only one estimate or could not be classified into pre-defined categories are listed in Table 2.6 overleaf, along with reported prevalence rates and specific reasons for exclusion from meta-analyses.

Table 2.6 Prevalence rates excluded from meta-analyses.

Study	Geographical location	Definition of joint problem	Reported prevalence	Construct	Reason for exclusion from meta-analyses
Depression:					
Dunlop, 2005	U.S.A	Arthritis	40.3%	Depression	Unclear cut-off point
Figaro,2005	U.S.A	ACR criteria	22.0%	2 questions: mood and anhedonia	Unclear classification of severity
Fisher, 2004	U.S.A	Arthritis	23.2%	Depression	Non-standardised tool
Jakobsson, 2006	Sweden	Self-reported OA	36.3%	Depression	Non-standardised tool
Kadam, 2007	UK	Medical records define OA	2.8%	Depressive disorders	Only one estimate
Mallen, 2008	UK	Symptomatic OA	18.2%	2 questions: mood and anhedonia	Unclear classification of severity
Moussavi, 2007	60 countries	Arthritis	Average: 10.7%	Depressive episode	Only one estimate
Muus, 2007	U.S.A	Arthritis	18.8%	Self-reported recall of doctor diagnosed depression	Unclear classification of severity
Polsky, 2005	U.S.A	Arthritis	4.4%	Depression	Unclear classification of severity
Szoeke, 2008	Australia	Self-reported OA	17.0%	Depression	Unclear cut-off point
Anxiety:					
Leveille, 2007	U.S.A	Symptomatic OA	19.0%	'Prevalent' anxiety	Unclear classification of severity
McWilliams, 2008	U.S.A	Arthritis	9.4%	Simple phobia	Only one estimate
Wells, 1989a	U.S.A	Arthritis	11.9%	Anxiety disorders	Only one estimate

Numbers of pooled estimates across pre-defined categories, and methods of ascertainment are displayed in Table 2.7 overleaf. Notably, prevalence rates of clinical interview defined depression and anxiety disorders are not only based on samples with 'arthritis', but the majority of prevalence rates have been derived from one publication (He et al., 2008).

Table 2.7 Numbers of pooled estimates and instruments used.

Construct	Number of prevalence estimates	Assessment instrument (number of prevalence rates estimated with the tool)
Depressive disorders		
Major depression	23	WMH WHO-CIDI (20)^, WMH WHO-CIDI-SF (2), AUDADIS (1)
Dysthymia	19	WMH WHO-CIDI (18)*, AUDADIS (1)
Depression symptoms		
'Mild or worse'	18	CESD-20 (7), HADS-D (6), GDS-15 (3), PHQ-9 (1),
'Moderate or worse'	10	GDS-30 (1)
		AIMS-D (2), HADS-D (6), CESD-20 (1), PHQ-9 (1)
Anxiety disorders		
GAD	19	WMH WHO-CIDI (18)*, AUDADIS (1)
Social phobia	19	WMH WHO-CIDI (18)†, AUDADIS (1)
Panic with agoraphobia	19	WMH WHO-CIDI (18)*, AUDADIS (1)
PTSD	18	WMH WHO-CIDI (18)*
Panic disorder	2	WMH WHO-CIDI (1), AUDADIS (1)
Anxiety symptoms		
'Mild or worse'	6	HADS-A (6)
'Moderate or worse'	5	HADS-A (5)

Note: AIMS-D - Arthritis Impact Measurement Scales-Depression; AUDADIS - Alcohol Use Disorders and Associated Disabilities Interview; CESD - Center for Epidemiologic Studies Depression Scale; GAD - generalised anxiety disorder; GDS - Geriatric Depression Scale; HADS-D, -A - Hospital Anxiety and Depression Scale-Depression and Anxiety subscale; PHQ - Patient Health Questionnaire; PTSD - post-traumatic stress disorder; WMH WHO-CIDI - World Mental Health (WMH) Survey Initiative version of the World Health Organization's Composite International Diagnostic Instrument;
Numbers of estimates derived from a study by He et al. (2008): * - all, ^ -18, † - 17.

2.6.4 Meta-analyses of questionnaire data for clinically relevant symptoms of depression and anxiety coexisting with osteoarthritis

2.6.4.1 Depression symptoms: questionnaire data

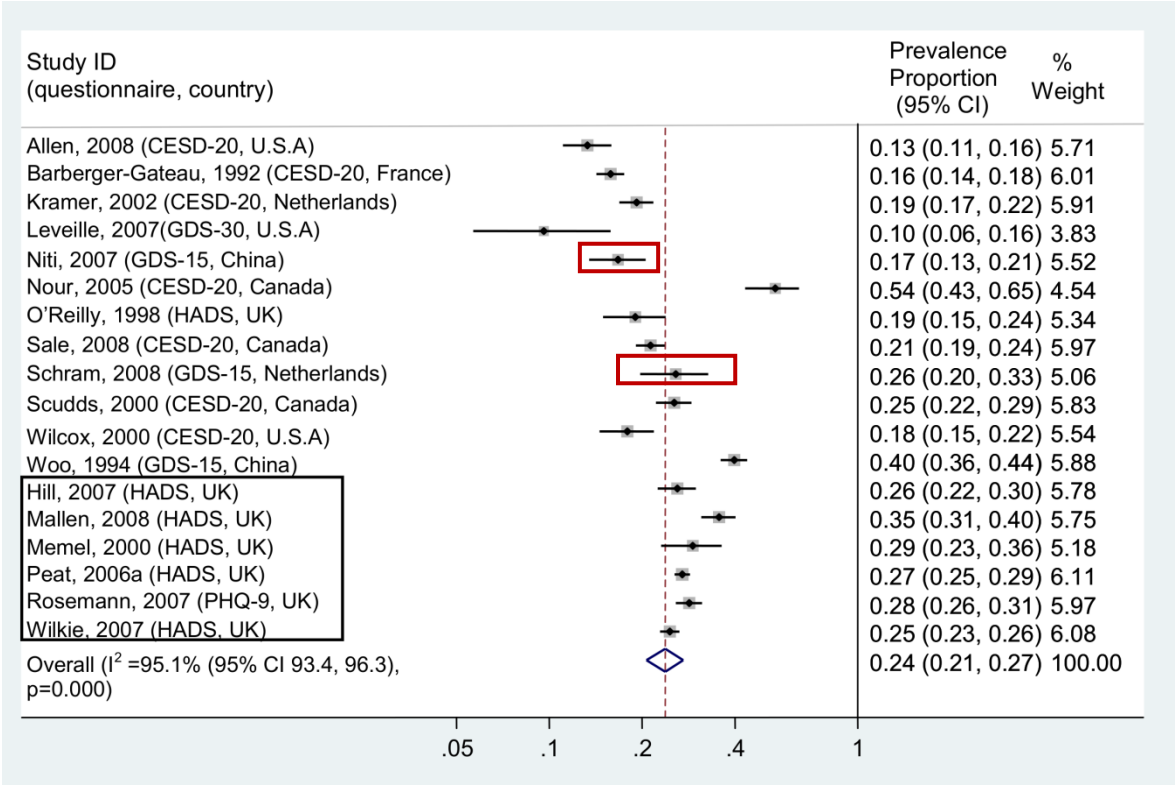
'Moderate or worse' severity of depression symptoms

Prevalence rates of 'moderate or worse' depression symptoms identified in individual studies ranged from 3.2% to 36.8%. A meta-analysis of ten prevalence estimates of 'moderate or worse' depression symptoms resulted in a pooled proportion of 14.6% (95% CI 9.9, 21.0) (Dexter & Brandt, 1994, Creamer et al., 1999, Memel et al., 2000, Kramer et al., 2002, Croft et al., 2005, Nour et al., 2005, Peat, 2006a, Rosemann et al., 2007, Wilkie et al., 2007, Mallen & Peat, 2008), with a significant and high inconsistency between studies (chi^2=387.4, p<0.0001, I^2=97.7% (95% CI 96.8, 98.3)) being evident (Table B.4.1. in Appendix B.4 on page 343), but no evidence of publication bias (p=0.395; see Table B.5.1 in Appendix B.5 on page 349).

'Mild or worse' severity of depression symptoms

Prevalence rates of 'mild or worse' symptoms ranged from 9.6% to 54.0%. Eighteen prevalence estimates of 'mild or worse' depression symptoms were meta-analysed (Barberger-Gateau et al., 1992, Woo et al., 1994, O'Reilly et al., 1998, Memel et al., 2000, Scudds et al., 2000, Wilcox et al., 2000, Kramer et al., 2002, Nour et al., 2005, Peat et al., 2006a, Hill et al., 2007, Leville et al., 2007, Niti et al., 2007, Rosemann et al., 2007, Wilkie et al., 2007, Allen et al., 2008, Mallen & Peat, 2008, Sale et al., 2008, Schram et al., 2008). A pooled proportion was 23.8% (95% CI 20.6, 27.2; see Figure 2.3 overleaf), with a significant and high level of heterogeneity (chi^2=344.89, p<0.0001, I^2=95.1% (95% CI 93.4, 96.3)) (Table B.4.2 in Appendix B.4 on page 343) and no evidence of publication bias (p=0.730; see Table B.5.1 in Appendix B.5 on page 349). Removal of two prevalence estimates from samples with 'arthritis' resulted in pooled prevalence of 24.2% (95% CI 20.8, 27.8), and still heterogeneity (chi^2=331.41, p<0.0001, I^2=95.5% (95% CI 93.9, 96.6)). This indicates unlikely impact on inclusion of samples defined as having 'arthritis' on pooled prevalence estimates.

Figure 2.3 Prevalence rates of 'mild or worse' depression symptoms in community-dwelling adults with osteoarthritis.



Note: CESD - Center for Epidemiologic Studies Depression Scale; GDS - Geriatric Depression Scale; HADS - Hospital Anxiety and Depression Scale-Depression subscale; PHQ - Patient Health Questionnaire;
Framed in black - primary care studies; Framed in red - two studies in adults with 'arthritis'.

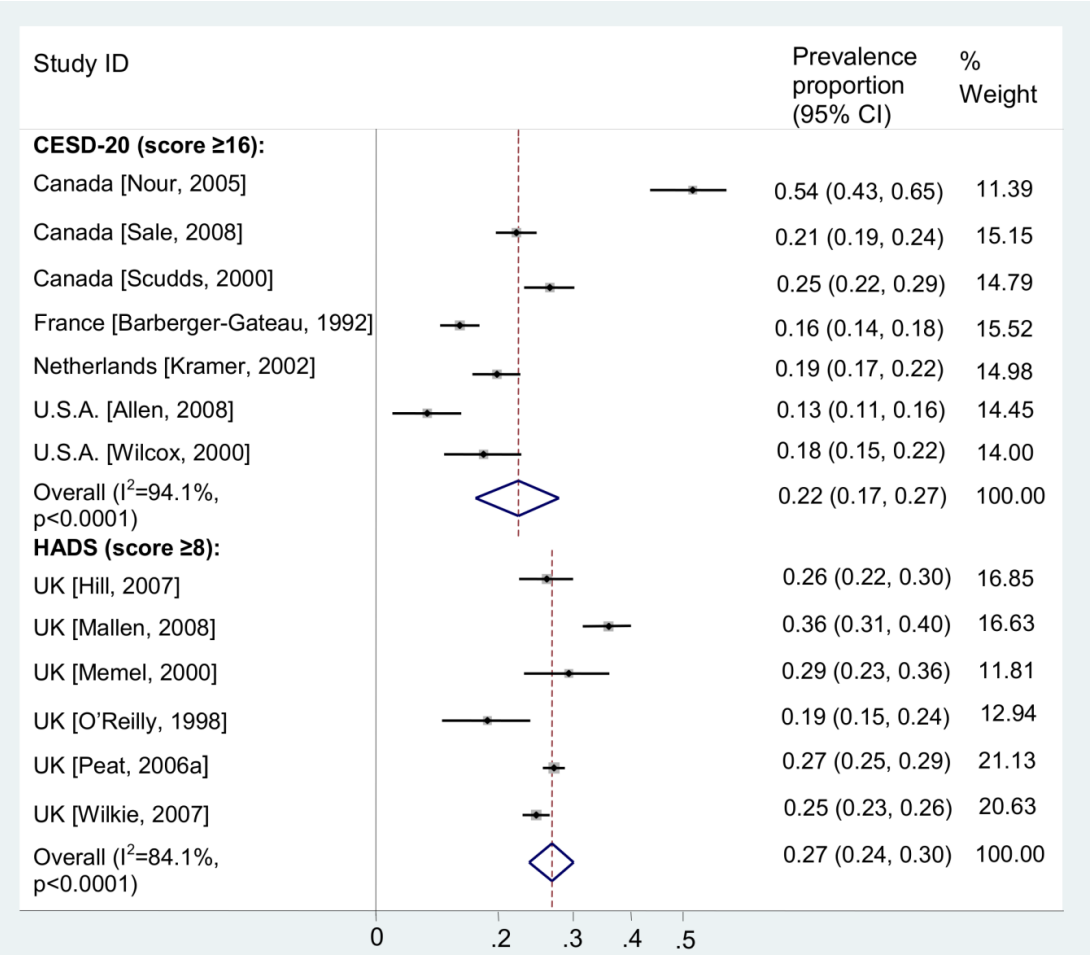
Pooled prevalence estimates by subgroups can be found in Table 2.8 (overleaf). The majority of pooled estimates ranged from 21% to 28%, with two outliers identified for the ACR criteria defined and medical records defined OA. Prevalence rates the two commonest questionnaires used to assess depression (CESD-20 and HADS-D) are displayed in Figure 2.4 (on page 52).

Table 2.8 Summary of subgroups of pooled estimates of 'mild or worse' depression symptoms arranged by % weight.

Grouping factor	No. of ES	ES. range	Pooled prevalence	95%CI	I^2	% weight
Questionnaire to assess depression symptoms:						
CESD-20 [1, 2, 11, 17, 22, 24, 26]	7	13.3,54.0	21.5	17.1,26.6	94.1	39.51
HADS-D [9, 13, 14,18, 19, 27]	6	19.0,35.5	26.7	23.7,30.0	84.1	34.24
GDS-15 [16, 23, 28]	3	16.7,39.8	26.5	14.0,44.2	96.8	16.46
PHQ-9 [21]	1	28.4	-	25.7,31.3	-	5.97
GDS-30 [12]	1	9.6	-	5.7,15.8	-	3.83
Definitions of OA:						
Symptomatic OA [2, 12, 13, 19, 22, 24, 27, 28]	8	9.6,39.8	24.4	19.8,29.6	96.6	45.46
Self-reported OA [9, 11]	2	19.2,26.0	22.4	16.4,29.7	89.5	11.68
Radiographic OA [1, 18]	2	13.3,19.0	22.9	14.2,34.8	81.7	11.51
ACR criteria [21, 26]	2	17.9,28.4	15.8	11.0,22.1	94.2	11.05
'Arthritis' [16, 23]	2	16.7,25.7	20.6	13.2,30.7	84.1	10.57
Medical records defined OA [14, 17]	2	29.2,54.0	40.8	19.8,65.7	93.0	9.73
Study setting:						
General Population [1, 2, 11, 12, 16, 17, 18, 22, 23, 24, 26, 28]	12	9.6,54.0	21.6	17.2,26.7	95.6	65.13
Primary Care [9, 13, 14, 19, 21, 27]	6	24.6,35.5	28.0	25.5,30.6	79.0	34.87
Geographical location:						
Europe [2, 9, 11, 13, 14, 18, 19, 21, 23, 27]	10	15.8,35,5	24.5	21.2,28.2	93.8	57.18
U.S.A, Canada [1, 12, 17, 22, 24, 26]	6	9.6,54.0	21.3	15.3,28.9	94.5	31.42
China [16, 28]	2	16.7,39.8	26.8	10.2,54.1	98.3	11.40
Sample size:						
>200 [1, 2, 9, 11, 13, 16, 18, 19, 21, 22, 24, 26-28]	14	15.8,39.8	23.8	19.7,24.5	95.6	81.39
≤200 [12, 14, 17, 23]	4	9.6,54.0	27.1	14.6,44.6	93.3	18.61
Age (years)* (range 61-79):						
>67 [2, 14, 17, 22, 24, 26]	6	15.8,54.0	25.1	19.2,32.1	94.5	33.07
≤67 [13, 18, 19, 21, 27]	5	19.0,35.5	26.9	23.7,30.3	87.6	29.25
Proportion of females*(52.4-75.6):						
>64 [13, 19, 24, 26, 27]	5	17.9,35.5	25.9	22.4,29.6	89.4	29.31
≤64 [1, 11, 14, 18, 21, 22, 28]	7	13.3,39.8	23.4	17.8,30.3	96.2	39.96
Study quality:						
≥Good [1, 9, 12, 13, 14, 16, 17, 18, 19, 21, 22, 24, 26, 27]	14	9.6,54.0	23.7	20.7,27.1	92.8	77.14
Poor/fair [2, 11, 23, 28]	4	15.8,39.8	24.1	14.7,36.8	98.0	22.86

Note: ACR - the American College of Rheumatology; CI - confidence intervals; CESD - Center for Epidemiologic Studies Depression Scale; ES -estimates; GDS - Geriatric Depression Scale; HADS - Hospital Anxiety and Depression Scale; OA - osteoarthritis; PHQ - Patient Health Questionnaire;
[a number in square brackets] - represents a study ID number that can be found in Table B.2.1 (in Appendix B.2 on page 328); * - Data partially unavailable.

Figure 2.4 Prevalence of 'mild or worse' depression symptoms assessed with CESD-20 and HADS in community-dwelling adults with osteoarthritis.



Note: CESD - Center for Epidemiologic Studies Depression Scale; HADS - Hospital Anxiety and Depression Scale.

2.6.4.2 Anxiety symptoms: questionnaire data

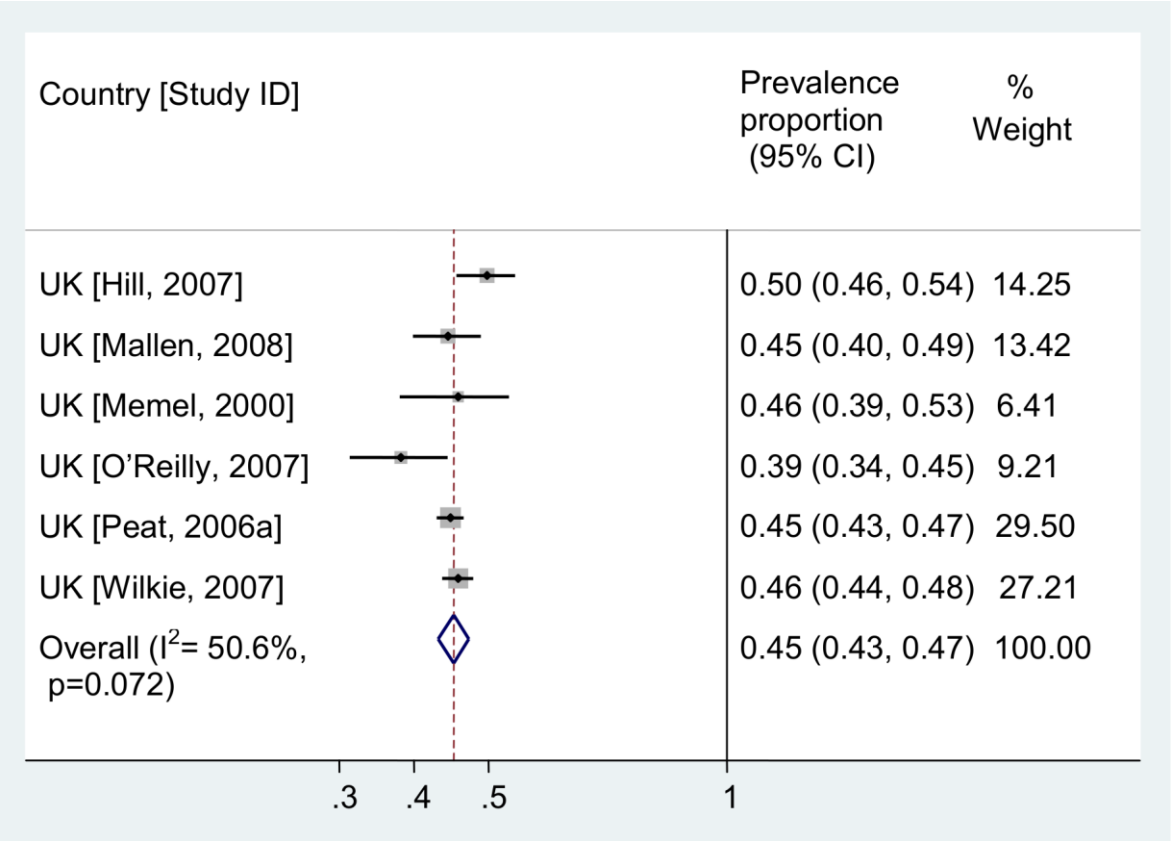
'Moderate or worse' levels of anxiety symptoms

Prevalence rates of 'moderate or worse' symptoms ranged across individual studies from 17.0% to 24.4%. Five prevalence estimates were meta-analysed (Memel et al., 2000, Croft et al., 2005, Peat et al., 2006a, Wilkie et al., 2007, Mallen & Peat, 2008). A pooled estimate was 20.8% (95% CI 18.0, 23.8; see Table B.4.3 in Appendix B.4, p. 344), with high heterogeneity (chi^2=23.91, p<0.0001, I^2=83.3% (95% CI 62.0, 92.6) and no evidence of publication bias (p=0.281; Table B.5.1 in Appendix B.5, p. 349).

'Mild or worse' levels of anxiety symptoms

Meta-analysed prevalence rates of 'mild or worse' symptoms of anxiety in patients with OA/ joint pain ranged from 39% to 50%. A pooled estimate was 45.4% (95% CI 43.4, 47.5; see Figure 2.5) with low heterogeneity (chi^2=10.12, p=0.072, I^2=50.6% (95% CI 0.0, 80.4)) (Table B.4.4 in Appendix B.4, p. 344) and no evidence of publication bias (p=0.854; see Table B.5.1, p. 349 for details).

Figure 2.5 Prevalence of 'mild or worse' anxiety symptoms in community-dwelling adults with osteoarthritis.



The low heterogeneity could be attributed to the consistency of the method of ascertainment, but also to the fact that five of the meta-analysed studies used a closely related sample-frame. Prevalence estimates for different sub-groups are presented in Table 2.9 overleaf.

Table 2.9 Summary of grouped estimates for 'mild or worse' symptoms of anxiety arranged by % weight.

Grouping factor	No. of ES	ES range	Pooled prevalence	95% CI	I^2	% weight
Definitions of OA:						
Symptomatic OA [13, 19, 27]	3	44.4,46.0	45.4	44.1,46.6	0.0	70.13
Self-reported OA [9]	1	50.0	-	46.0, 54.2	-	14.25
Radiographic OA [18]	1	39.0	-	33.6, 44.6	-	9.21
Medical records defined OA [14]	1	46.0	-	39.0, 53.0	-	6.41
Study setting:						
Primary care [9, 13, 14, 19, 27]	5	44.4,50.0	45.9	44.4, 47.3	17.8	90.79
General population [18]	1	39.0	-	33.6, 44.6	-	9.21
Sample size:						
>200 [9, 13, 18, 19, 27]	5	39.0,50.0	45.4	43.2, 47.6	60.4	93.59
≤200 [14]	1	46.0	-	38.9, 53.3	-	6.41
Mean age (years)* (range 61-71):						
>65 [14, 19, 27]	3	45.0,46.0	45.4	44.1, 46.8	0.0	63.11
≤65 [13, 18]	2	39.0,44.4	42.1	36.7, 47.8	59.6	22.63
Proportion of females* (range 58-65):						
>60 [19, 27]	2	45.0,46.0	45.4	44.1, 46.6	0.0	56.71
≤60 [13, 14, 18]	3	39.0,46.0	43.1	39.0, 47.2	37.2	29.04
Unfeasible for sub-group analyses:						
Questionnaire to assess anxiety symptoms:	consistent (HADS-A)					
Geographical location:	consistent (UK)					
Study quality:	consistent ('≥ good')					

Note: CI - confidence intervals; ES - estimates;
[a number in square brackets] - represents a study ID number that can be found in Table B.2.1 (in Appendix B.2 on page 328); ***** - Data partially unavailable.

2.6.5 Meta-analyses of clinical interview data in people with arthritis

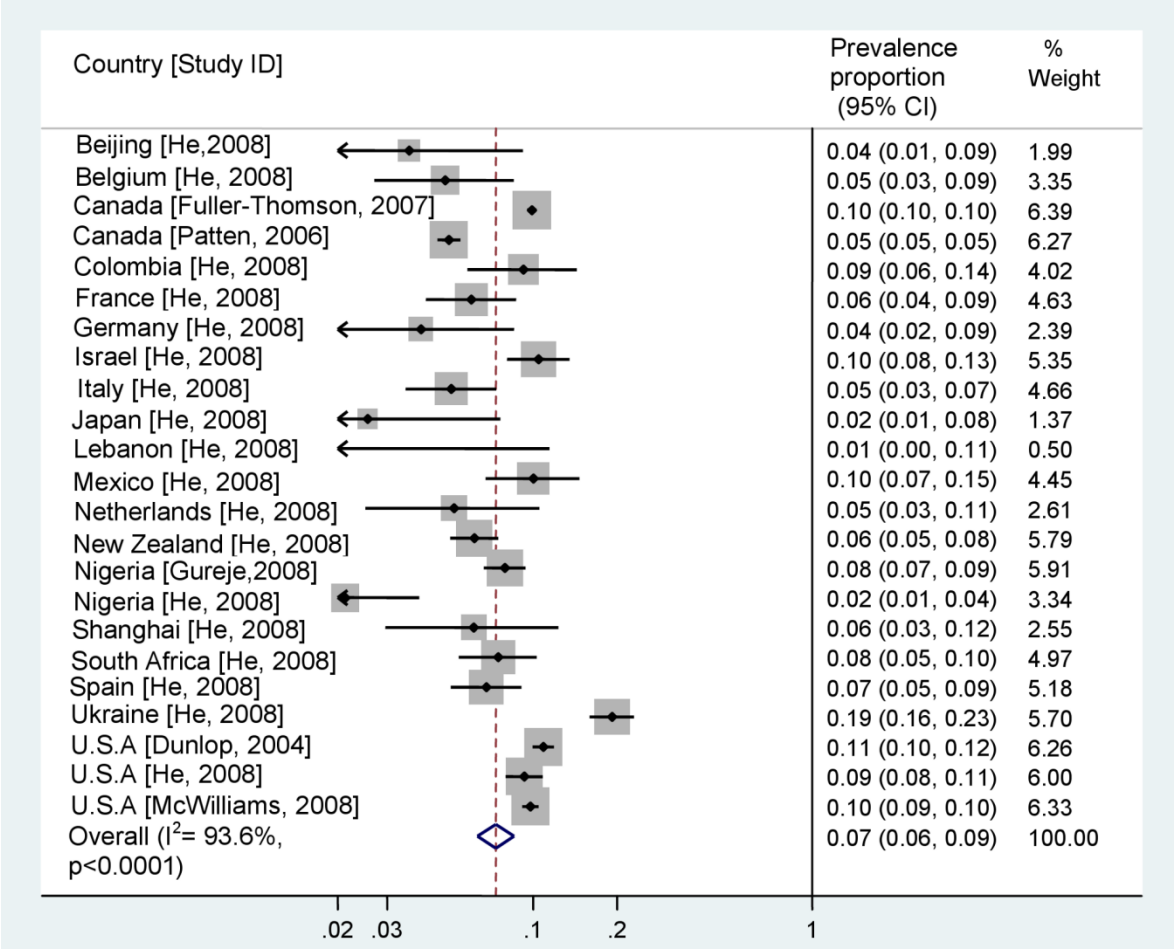
2.6.5.1 Depressive disorders

Major depression (MD)

Prevalence rates of major depression reported in individual studies ranged from 1.4% to 19.2%. A meta-analysis was conducted for 23 12-month prevalence estimates of major depression (MD) in patients with OA/ joint pain (Dunlop et al., 2004, Patten et al., 2006, He et al., 2008, Gureje et al., 2008, McWilliams et al., 2008, Fuller-Thomson & Shaked, 2009). This

meta-analysis revealed a 7.3% (95% CI 6.3, 8.5; see Figure 2.6) pooled prevalence rate of major depression, high inconsistency between studies (chi^2=342.54, p<0.0001, I^2=93.6% (95% CI 91.6, 95.1)) (see Table B.4.5 in Appendix B.4, p. 345), and no evidence of publication bias (p=0.071; see Table B.5.2 in Appendix B.5, p. 350).

Figure 2.6 Prevalence of major depression in community-dwelling adults with 'arthritis'.

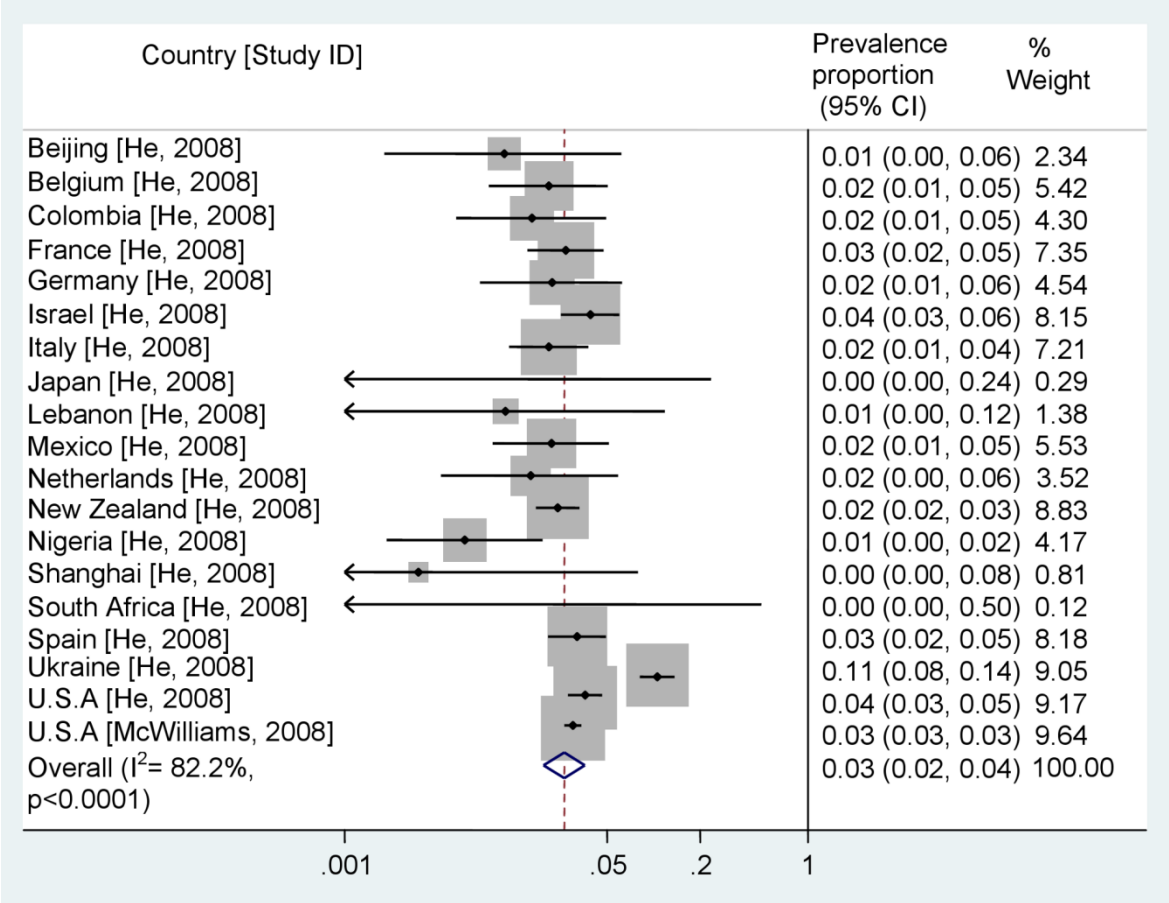


Dysthymic disorder (dysthymia)

Prevalence rates of dysthymia reported in individual studies (including 18 from one multinational study) ranged from 0.0% to 10.6%. Nineteen 12-month prevalence estimates of dysthymia were meta-analysed (He et al., 2008, McWilliams et al.,

2008)]. The pooled prevalence rate was 2.6 % (95% CI 2.0, 3.6; see Figure 2.7), with high heterogeneity (chi^2=100.98, p<0.0001, I^2=82.2% (95% CI 73.2, 88.1)) (see Table B.4.6 in Appendix B.4 on page 345), and no evidence of publication bias (p=0.155; see Table B.5.2 in Appendix B.5 on page 350 for details).

Figure 2.7 Prevalence of dysthymia in community-dwelling adults with 'arthritis'.



Sub-group analyses for major depression and dysthymia in community based adults with 'arthritis'

Sub-group meta-analyses were feasible only for geographical location, sample size and assessment tool. The results for major depression and dysthymia are displayed in Table 2.10 overleaf. Details of variables unfeasible for sub-group analyses are also presented this table.

Table 2.10 Summary of subgroups of pooled estimates of major depression and dysthymia arranged by % weight.

Grouping factor	No. of ES	ES range	Pooled prevalence	95% CI	I^2	% weight
Major Depression (MD)						
Geographical location:						
U.S.A, Canada [29, 31, 33, 35, 36]	5	5.0,10.9	8.7	6.9,11.00	98.0	31.24
West and South Europe [31]	6	3.8,6.8	5.7	4.8, 6.8	0.0	22.83
Africa [30, 31]	3	2.2,7.9	5.6	3.2,9.5	88.0	14.22
Colombia, Mexico [31]	2	9.3,10.2	9.8	7.3, 13.0	0.0	8.46
East Asia [31]	3	2.2,5.7	4.2	2.4,7.1	0.0	5.91
Middle East [31]	2	1.4,10.5	5.2	0.8, 28.6	71.0	5.85
New Zealand [31]	1	6.2	-	5.0, 7.6	-	5.71
Ukraine [31]	1	19.2	-	15.9, 23.0	-	5.70
Sample size:						
> 200 [29, 30, 31, 33, 35, 36]	18	2.2,19.2	7.8	6.7,9.3	95.3	84.57
≤ 200 [31]	7	1.4,9.2	4.9	3.3,7.4	41.3	15.43
Assessment tool:						
WMH WHO-CIDI [30, 31, 36]	20	1.4,19.2	6.5	5.2,8.1	90.9	81.03
WMH WHO-CIDI-SF [29, 35]	2	9.9,10.9	10.3	9.4,11.2	71.9	12.65
AUDADIS [33]	1	9.8	-	9.2,10.5	-	6.33
Dysthymia:						
Geographical location:						
West and South Europe [31]	6	1.6,3.2	2.6	2.0, 3.4	0.0	36.21
U.S.A [31, 33]	2	3.0,3.6	3.2	2.7, 3.8	36.6	18.81
Colombia, Mexico [31]	2	1.6,2.0	2.0	1.0, 3.8	0.0	9.83
Middle East [31]	2	1.1,3.9	3.8	2.3, 5.8	0.5	9.53
Ukraine [31]	1	10.6	-	8.2, 13.7	-	9.05
New Zealand [31]	1	2.4	-	1.8, 3.3	-	8.83
Africa [31]	2	0.0,0.6	0.6	0.2, 2.0	0.0	4.29
East Asia [31]	3	0.1,1.1	0.7	0.2, 3.2	0.0	3.45
Sample size:						
>200 [31, 33]	12	0.0,10.6	3.0	2.2,4.2	88.0	82.81
≤200 [31]	7	0.1,2.2	1.5	0.8,2.7	0.0	17.19
Assessment tool:						
WMH WHO-CIDI [31]	18	0.0,10.6	2.4	1.7,3.6	82.2	90.36
AUDADIS [33]	1	3.0	-	2.6,3.4	-	9.64
Unfeasible for sub-group analyses:						
Definition of OA:	consistent (‘arthritis’)					
Study setting:	consistent (general population)					
Proportion of females:	>50% data unavailable					
Study quality:	consistent (poor/fair)					
Mean age:	>50% data unavailable					

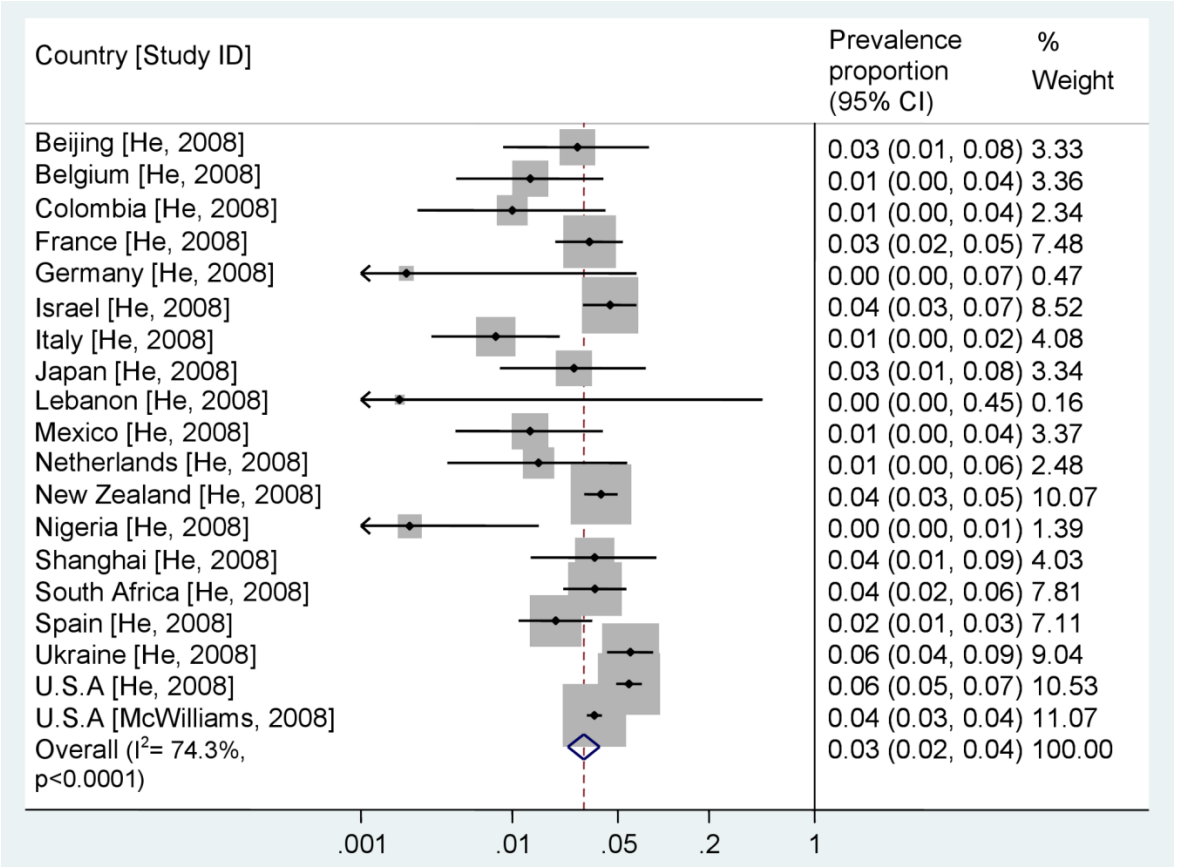
Note: AUDADIS- Alcohol Use Disorders and Associated Disabilities Interview; CI- confidence intervals; ES- estimates; WMH WHO-CIDI (SF) - World Mental Health (WMH) Survey Initiative version of the World Health Organization's Composite International Diagnostic Instrument (short version).
[a number in square brackets] - represents a study ID number that can be found in Table B.2.1 (in Appendix B.2 on page 328).

2.6.5.2 Anxiety disorders

Generalised anxiety disorder (GAD)

A meta-analysis of 19 prevalence rates (range 0.2% to 6.0%) [(He et al., 2008, McWilliams et al., 2008)] of 12-month prevalence of generalised anxiety disorder (GAD) produced a pooled estimate of 3.0% (95% CI 2.5, 3.9; see Figure 2.8), with moderate heterogeneity (chi^2=69.91, p<0.0001, I^2=74.3% (95% CI 59.6, 83.6)) (see Table B.4.7 in Appendix B.4, p. 346) and evidence of statistically significant publication bias (p=0.017; see Table B.5.2 in Appendix B.5, p. 350). This suggests that precision in estimating pooled prevalence of GAD increases as the sample size of component studies increases.

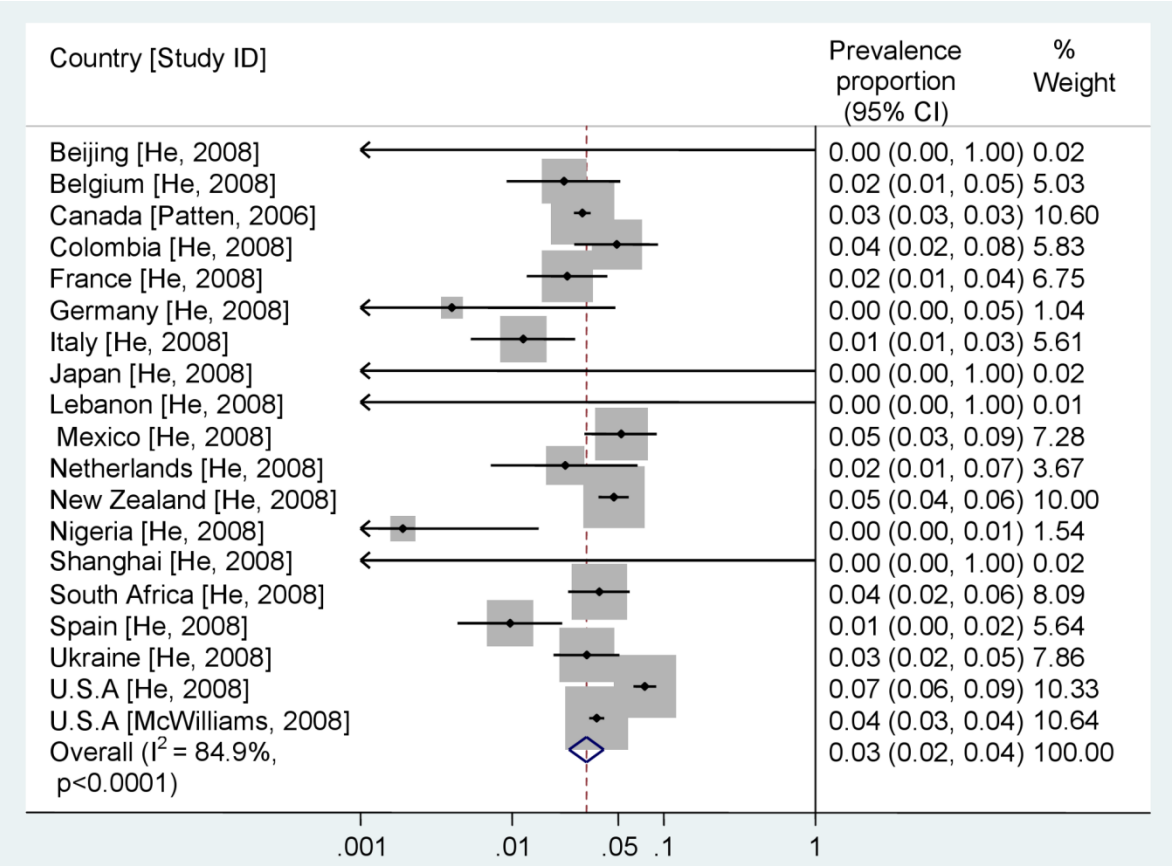
Figure 2.8 Prevalence of GAD in community-dwelling adults with 'arthritis'.



Social phobia

A meta-analysis of 19 12-month prevalence rates (range 0.0% to 7.5%) of social phobia [(He et al., 2008, McWilliams et al., 2008)] revealed a 3.0% pooled prevalence estimate (95% CI 2.3, 3.9; see Figure 2.9), with no evidence of publication bias (p=0.198; see Table B.5.2 in Appendix B.5 on page 350) and a high level of heterogeneity (chi^2=118.84, p<0.0001, I^2=84.9% (95% CI 77.6, 89.7)) (see Table B.4.8 in Appendix B.4 on page 346).

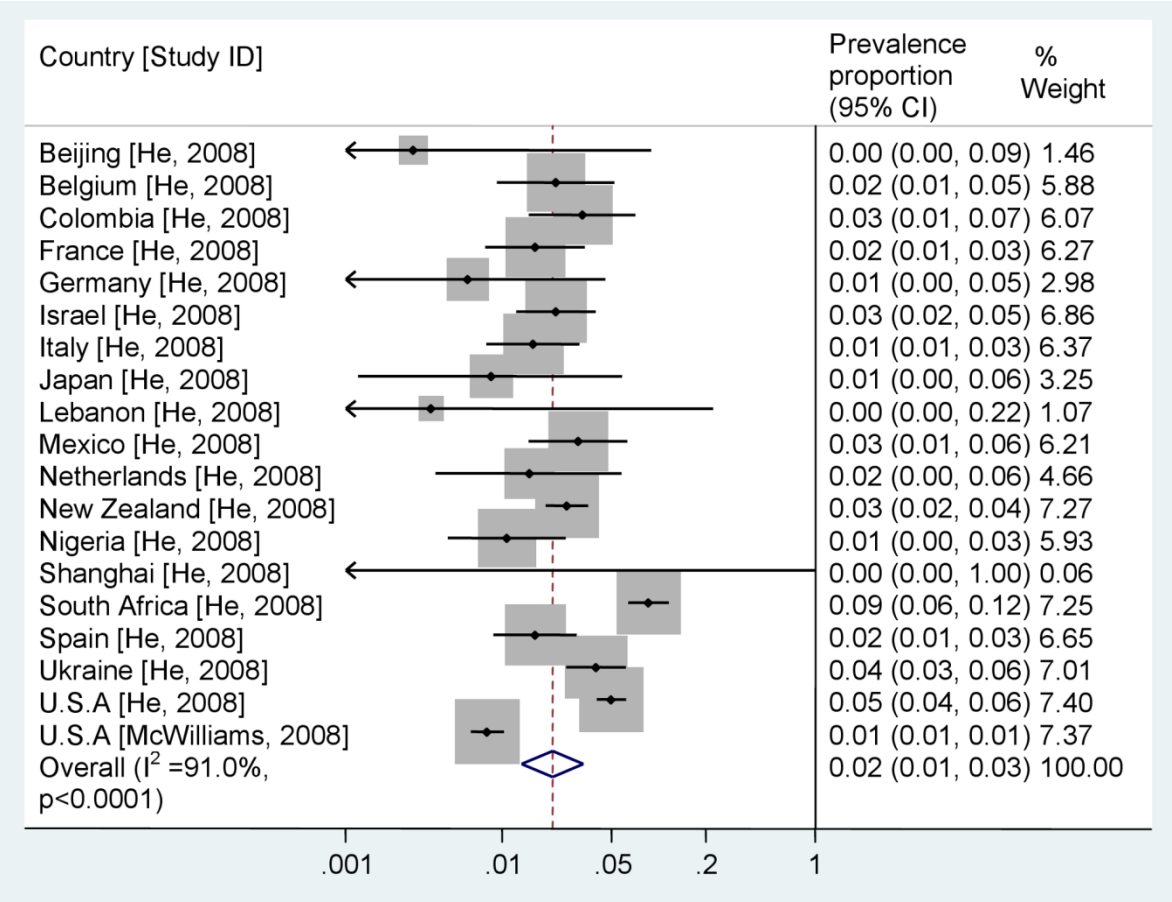
Figure 2.9 Prevalence of social phobia in community-dwelling adults with 'arthritis'.



Panic with agoraphobia

A meta-analysis of 19 12-month prevalence rates (range 0.0% to 8.5%) of panic with agoraphobia (He et al., 2008, McWilliams et al., 2008) found a 2.2% pooled prevalence rate (95% CI 1.4, 3.4; see Figure 2.10), with no evidence of publication bias (p=0.409; see Table B.5.2 in Appendix B.5 on page 350) and high heterogeneity (chi^2=199.35, p<0.0001, I^2=91.0% (95% CI 87.4, 93.5)) (see Table B.4.9 in Appendix B.4 on page 347).

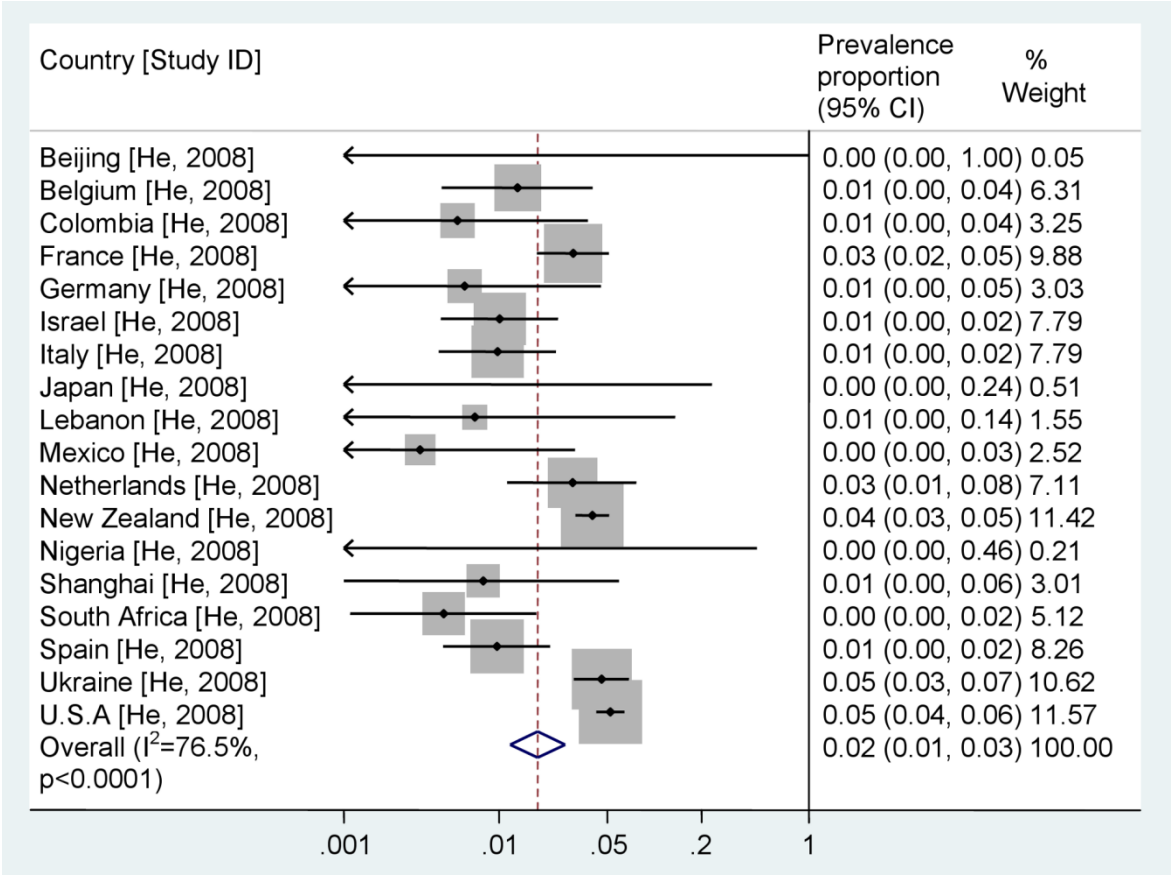
Figure 2.10 Prevalence of panic disorder with agoraphobia in community-dwelling adults with 'arthritis'.



Post-traumatic stress disorder (PTSD)

Eighteen 12 month prevalence rates included in meta-analyses ranged from 0.0% to 5.2% [(He et al., 2008)]. A pooled estimate of prevalence PTSD was 1.8% (95% CI 1.2, 2.6, see Figure 2.11) with high heterogeneity (chi^2=72.25, p<0.0001, I^2=76.5% (95% CI 63.0, 85.0)) (see Table B.4.10 in Appendix B.4, p. 347). Statistically significant publication bias was found with the Egger's test (p<0.0001; see Table B.5.2 in Appendix B.5, p. 350), what is likely to be due to 10 estimates being 0% or close to 0%.

Figure 2.11 Prevalence of PTSD in community-dwelling adults with 'arthritis'.



Panic disorder

A meta-analysis of 2 (2.6% and 3.0%) 12 month prevalence rates [Patten et al., 2006, McWilliams et al., 2008] of panic disorder resulted in a pooled estimate of 2.8% (95% CI 2.4, 3.2; see Table B.4.11 in Appendix B.4 on page 348), with high levels of heterogeneity detected (chi^2=82.60, p<0.0001, I^2=98.8% (95% CI 97.5, 99.4)). Whilst by definition two estimates were sufficient to perform meta-analysis, the STATA programme did not allow exploring a publication bias analysis for two estimates.

Sub-group analyses for specific anxiety disorders

Sub-group estimates were unattainable for panic disorder, as only 2 prevalence rates were available. Data for other specific anxiety disorders was derived from three studies [Patten et al., 2006, He et al., 2008, McWilliams et al., 2008], and thus, there was a limited variability across the majority of pre-defined variables of interest. As the identified studies used cross-sectional study design and reported results for several countries with limited space for details about characteristics of sub-samples, problems of consistency across studies or unavailable of data were common. As a result, analyses could be performed for geographical location (see Table 2.11 overleaf) and sample size and assessment tool only (see Table 2.12 on page 64). Details of variables unfeasible for sub-group analyses are presented in Table 2.12.

Table 2.11 Summary of estimates grouped by geographical locations across anxiety disorders arranged by % weight.

Geographical location	No. of ES	ES. range	Pooled prevalence	95% CI	I^2	% weight
GAD:						
West and South Europe [31]	6	0.2,3.2	1.7	1.0,2.8	46.0	24.99
U.S.A [31, 33]	2	3.5,5.9	4.5	2.7,7.5	95.1	21.60
East Asia [31]	3	2.6,3.6	2.9	1.6,5.4	0.0	10.70
New Zealand [31]	1	3.9	-	3.0,5.0	-	10.07
Ukraine [31]	1	6.0	-	4.2,8.5	-	9.04
Africa [31]	2	0.2,3.5	1.0	0.0,14.3	86.7	9.20
Middle East [31]	2	0.2,4.5	3.9	1.2,12.4	7.1	8.69
Colombia, Mexico [31]	2	1.0,1.4	1.2	3.2,6.5	0.0	5.71
Social phobia:						
U.S.A, Canada [31, 33, 36]	3	2.9,7.5	4.3	2.6,6.9	97.4	31.58
West and South Europe [31]	6	0.4,2.2	1.6	1.1,2.2	0.0	27.74
Colombia, Mexico [31]	2	3.7,5.3	4.7	3.0,7.1	0.0	13.11
New Zealand [31]	1	4.7	-	3.8,5.9	-	10.00
Africa [31]	2	0.2,3.8	1.1	0.0,16.1	87.3	9.64
Ukraine [31]	1	3.1	-	2.0,5.2	-	7.86
Lebanon [31]	1	0.0	-	0.0,100.0	-	0.01
East Asia [31]	3	0.0	0.0	0.0,80.0	0.0	0.07
Panic with agoraphobia:						
West and South Europe [31]	6	0.6,2.2	1.7	1.2,2.3	0.0	32.80
U.S.A [31, 33]	2	0.8,5.0	2.1	0.0,11.5	99.2	14.77
Africa [31]	2	1.1,8.5	3.2	0.0,21.0	95.0	13.18
Colombia, Mexico [31]	2	2.9,3.2	3.0	1.8,5.2	0.0	12.28
U.S.A [31]	1	5.2	-	4.2,6.5	-	11.57
Middle East [31]	2	0.4,2.9	2.8	1.7,4.7	0.0	7.92
New Zealand [31]	1	2.6	-	1.9,3.6	-	7.27
Ukraine [31]	1	4.0	-	2.5,6.2	-	7.01
East Asia [31]	3	0.0,0.9	0.7	0.0,3.4	0.0	4.77
PTSD:						
West and South Europe [31]	6	0.6,3.1	1.6	1.0,2.7	50.8	42.38
U.S.A [31]	1	5.2	-	4.2,6.5	-	11.57
New Zealand [31]	1	4.0	-	3.1,5.1	-	11.42
Ukraine [31]	1	4.5	-	3.0,6.9	-	10.62
Middle East [31]	2	0.7,1.1	1.0	0.4,2.2	0.0	9.35
Colombia, Mexico [31]	2	0.3,0.6	0.4	0.0,1.9	0.0	5.77
Africa [31]	2	0.0,0.5	0.4	0.0,1.6	0.0	5.33
East Asia [31]	3	0.0,0.8	0.6	0.0,4.0	0.0	3.57

Note: CI- confidence intervals; GAD- generalise anxiety disorder; ES-estimates; PTSD- post traumatic stress disorder;
[a number in square brackets] - represents a study ID number that can be found in Table B.2.1 (in Appendix B.2 on page 328).

Table 2.12 Summary of estimates grouped by sample size and assessment tool across anxiety disorders.

Grouping factor	No. of ES	ES. range	Pooled %	95% CI	I^2	% weight
GAD:						
Sample size:						
>200 [31, 33]	12	0.2,6.0	3.2	2.4,4.1	81.7	83.84
≤200 [31]	7	0.2,3.6	2.2	1.3,3.6	0.0	16.16
Assessment tool:						
WMH WHO-CIDI [31]	18	0.2,6.0	2.7	2.1,3.7	74.1	88.93
AUDADIS [33]	1	3.5	-	3.1,3.9	-	10.07
Social phobia:						
Sample size:						
>200 [31, 33, 36]	12	0.2,7.5	3.1	2.3,4.1	90.2	89.39
≤200 [31]	7	0.0,3.7	2.7	1.5,4.9	0.0	10.61
Assessment tool:						
WMH WHO-CIDI [31, 36]	18	0.0,7.5	2.8	2.0,3.9	85.6	89.36
AUDADIS [33]	1	3.6	-	3.6,4.0	-	10.64
Panic with agoraphobia:						
Sample size:						
>200 (31, 33)	12	0.8,8.5	2.4	1.5,4.0	94.3	80.45
≤200 (31)	7	0.0,3.2	1.9	0.9,3.4	0.0	19.35
Assessment tool:						
WMH WHO-CIDI [31]	18	0.0,8.5	2.5	1.8,3.5	78.5	92.63
AUDADIS [33]	1	0.8	-	0.6,1.0	-	7.37
PTSD:						
Sample size:						
>200 [31]	11	0.0, 5.2	2.0	1.3, 3.1	83.3	81.49
≤200 [31]	7	0.0, 3.1	1.4	0.7, 2.9	0.0	18.51
Assessment tool:	consistent (WMH WHO-CIDI)					
Unfeasible for sub-group analyses:						
Definition of OA:	consistent (‘arthritis’)					
Study setting:	consistent (general population)					
Proportion of females:	> 50% data unavailable					
Study quality:	consistent (poor/fair)					
Mean age:	> 50% data unavailable					

Note: AUDADIS- Alcohol Use Disorders and Associated Disabilities Interview; CI- confidence intervals; ES- estimates; WMH WHO-CIDI- World Mental Health (WMH) Survey Initiative version of the World Health Organization's Composite International Diagnostic Instrument;
[a number in square brackets]- represents a study ID number that can be found in Table B.2.1 (in Appendix B.2 on page 328).

2.6.6 Meta-regression analyses of prevalence estimates of depressive/ anxiety disorders and symptoms

Meta-regression analyses found that the 10% level required for adjusted analyses was achieved only for covariates associated with prevalence rates of major depression in an unadjusted model, but failed to reach the pre-defined level of statistical significance in an adjusted model. Details of unadjusted models estimates for questionnaire data, with specified reason for omitting are displayed in Table 2.13 overleaf. Results of unadjusted models of depressive and anxiety disorders and the adjusted model for major depression ascertain with clinical interview schedules and reasons for omitting specific covariates are listed in Table 2.14 on page 67.

Table 2.13 Meta-regression analyses for prevalence estimates of questionnaire data, values show β, se (β), and the significance of β.

Construct	Model	Covariate							
		Age	**Sex**	**Setting**	**Geographical location**	**Sample size**	**OA def.**	**Tool**	**Study quality**
		Mean age	**Proportion of females**	**General population vs. Primary care**	**America vs. other**	**<200 vs. ≤200**	**Robust vs. unclear**	**CESD vs. other** **HADS vs. other**	**Fair/poor vs. good**
Depression symptoms									
'Mild or worse'	Unadj.	0.01 (0.03) p=0.704	-0.22 (2.18) p=0.921	0.36 (0.26) p=0.183	0.21 (0.27) p=0.463	0.24 (0.32) p=0.464	0.24 (0.27) p=0.385	0.18 (0.26) p=0.511	Omitted (100% good)
'Moderate or worse'	Unadj.	0.03 (0.05) p=0.486	-2.32 (1.69) p=0.208	0.58 (0.47) p=0.257	Omitted (70% Europe)	-0.19 (0.55) p=0.738	Omitted (70% robust)	0.69 (0.44) p=0.156	Omitted (100% good)
Anxiety symptoms									
'Mild or worse'	Unadj.	0.03 (0.02) p=0.234	-2.52 (1.98) p=0.209	Omitted (83% primary care)	Omitted (100% Europe)	Omitted (83% ≤200)	Omitted (100% robust)	Omitted (100% HADS)	Omitted (100% good)
'Moderate or worse'	Unadj.	Omitted (80% 65 years)	-0.39 (1.91) p=0.851	Omitted (100% primary care)	Omitted (100% Europe)	Omitted (80% ≤200)	Omitted (100% robust)	Omitted (100% HADS)	Omitted (100% good)

Note: Omitted - for reasons described in result section concerning sub-group meta-analyses; OA – osteoarthritis; Unadj.- unadjusted.

Table 2.14 Meta-regression analyses for prevalence estimates of clinical interview data, values show β, se (β), and the significance of β.

Construct	Model	Covariate		
		Geographical location	Sample size	Assessment tool
		America vs. other	>200 vs. ≤ 200	WMH WHO-CIDI vs. other
Depressive disorders				
Dysthymia	Unadj.	-0.07(0.39) p=0.859	0.76 (0.43) p=0.910	Omitted (95% assessed with WMH WHO-CIDI)
Major depression	Unadj.	-0.41(0.22) p=0.081	0.51 (0.27) p=0.078	Omitted (87% assessed with WMH WHO-CIDI)
	Adj.	-0.35(0.22) p=0.122†	0.44 (0.27) p=0.117‡	-
Specific anxiety disorders				
GAD	Unadj.	-0.17(0.42) p=0.694	0.42 (0.43) p=0.341	Omitted (95% assessed with WMH WHO-CIDI)
Panic with agoraphobia	Unadj.	-0.12 (0.43) p=0.783	0.58 (0.51) p=0.273	Omitted (95% assessed with WMH WHO-CIDI)
PTSD	Unadj.	-0.17(0.68) p=0.801	0.61 (0.61) p=0.339	Omitted (100% assessed with WMH WHO-CIDI)
Social phobia	Unadj.	-0.72 (0.10) p<0.05	0.29 (0.54) p=0.569	Omitted (95% assessed with WMH WHO-CIDI)

Note: Adj.- adjusted; Unadj.- unadjusted; GAD- generalised anxiety disorder; PTSD- Post Traumatic Stress Disorder; WMH WHO-CIDI- World Mental Health (WMH) Survey Initiative version of the World Health Organization's Composite International Diagnostic Instrument;
†- adjusted for sample; ‡- adjusted for geographical location;
Analyses for panic disorder and other covariates where omitted, for reasons described in result section concerning sub-group meta-analyses.

2.6.7 Supplementary analyses

The impact of the individual studies on the heterogeneity of prevalence rates were explored using the non-standardised and sequential algorithm methods. Possibly due to the random-effects approach, the non-standardised method involving 95% CI was of limited effectiveness, as the majority of the estimates lay beyond the combined 95% CI. There was no clear superiority for the sequential algorithm method. Both methods resulted in high proportions of studies being removed and can be regarded of limited utility to meta-analyses of prevalence rates with a random effect approach (see Table B.6.1 in Appendix B.6 on page 351 for a summary).

2.6.8 Summary of findings for the prevalence of anxiety and depression

Clinically relevant symptoms of depression and anxiety are common among people with OA. This systematic review and meta-analysis has found that 21.0% and 15.0% of people with OA have anxiety and depression symptoms classed as being ‘moderately severe or worse’ respectively. When mild symptoms are included, these estimates increase to 45.0% and 24.0% respectively. With the exception of major depression (7.3%), specific comorbid anxiety and depressive disorders in people with OA/ joint pain are comparatively rare (<5.0%). Between-study heterogeneity of estimates was typically high ranging from 50.6% to 98.8%. In total, 13 studies prevalence estimates could not be included in meta-analyses due to lack of specification of the used cut-off points, unclear severity of symptoms or conceptualisation that could not be classified as any of the used constructs. A summary of details of pooled prevalence estimates, included estimates and meta-regression analyses can be found displayed in Table 2.15 (on page 70).

Sub-group estimates suggest variation in prevalence of depression and anxiety in community-dwelling adults with OA across different sampling procedures, characteristic of sample and different methods of ascertainment. Marked differences appear, particularly for different geographical locations and for the method of determining anxiety and depression. Meta-regression analyses revealed the explored characteristics at the study level had small, inconsistent or statically non-significant effects of heterogeneity of pooled estimates.

Table 2.15 Summary of pooled prevalence estimates.

Construct	No. of ES	Total combined study population	Prevalence % rates range	Pooled % estimate (95% CI)	I^2 (95% CI)	Range of quality scores of analysed studies	Publication bias	Statistically significant sources of heterogeneity
Questionnaire assessed symptoms:								
'Moderate or worse' depression symptoms	10	9,005	3.2,36.8	14.6 (9.9,21.0)	97.7 (96.8,98.3)	10,14	No	No
'Mild or worse' depression symptoms	18	15,194	9.6,54.0	23.8 (20.6,27.2)	95.1 (93.4,96.3)	8, 15	No	No
'Moderate or worse' anxiety symptoms	5	6,489	17.0,24.4	20.8 (18.0,23.8)	83.3 (62.0,92.,6)	11,13	No	No
'Mild or worse' anxiety symptoms	6	6,867	39.0,50.0	45.4 (43.4,47.5)	50.6 (0.0,80.4)	11,15	No	No
Clinical interviews defined disorders:								
Dysthymia	19	15,718	0.0,10.6	2.6 (2.0,3.6)	82.2 (73.2,88.1)	9,10	No	No
Generalised anxiety disorder	19	15,718	0.2,6.0	3.0 (2.5,3.9)	74.3 (59.6,83.6)	9,10	Yes	No
Major depression	23	52,768	1.4,19.2	7.3 (6.3,8.5)	93.6 (91.6,95.1)	6,10	No	No
Panic disorder	2	16,121	2.6,3.0	2.8 (2.4,3.2)	98.8 (97.5,99.4)	8,9	-	No
Panic with agoraphobia	19	15,718	0.0,8.5	2.2 (1.4,3.4)	91.0 (87.4,93.5)	9,10	No	No
Post-traumatic stress disorder	18	7,842	0.0,5.2	1.8 (1.2,2.6)	76.5 (63.0,85.0)	10	Yes	No
Social Phobia	19	23,467	0.0,7.5	3.0 (2.3,3.9)	84.9 (77.6,89.7)	8,10	No	Study location in America

Note: ES – estimate; CI - confidence intervals.

2.7 DISCUSSION

2.7.1 Summary of key findings

This review suggests that symptoms of possible and probable depression and anxiety are common among community-dwelling adults with OA/joint pain. Analyses of samples of patients with an arthritis diagnosis, suggest that in total major depression (7%) and dysthymia (3%) can affect around 10% of people with osteoarthritis whilst specific anxiety disorders are rarely coexisting with osteoarthritis. Despite being more prevalent, symptoms of anxiety have received limited research attention in comparison to depression. In accordance with evidence derived from other populations of people with painful conditions (Routledge et al., 2006, Mitchell et al., 2011) and older adults (Luppa et al., 2010), heterogeneity of prevalence estimates was typically high. This appeared, in part, to be attributable to geographical location and methods of ascertaining anxiety and depression. Variance in pooled estimates could not be clearly accounted for between-study variance. Together, these observations suggest that depression and anxiety symptoms vary across different groups of people with osteoarthritis, and that this variability may be related to factors that could not be explored in this review, such as the severity of joint pain or the patients' economic situation.

2.7.2 Sources of possible bias in the reviewed evidence

The following section will examine sources of publication, selection and information bias in the reviewed evidence.

Publication bias

Publication bias can be defined as the tendency for articles containing 'positive or new findings' being more likely to be published (Porta et al., 2008). To reduce the risk of this type of bias, unpublished literature was searched using Web of Science and CSI Illumina. Unfortunately, identified unpublished evidence included dissertations that were difficult to access or conference publications that failed to meet the full-text inclusion criterion. However, a search for full-text publications (based on potentially relevant retrieved dissertations or conference publications) was conducted. This method appeared to be of limited use, as only a couple of full-text articles could be retrieved, both of which had been already identified through the systematic search.

Publication bias also refers to the tendency of authors to selectively report results that conform to their predefined notions (Porta et al., 2008). Indeed, ten studies could not be included, as the authors failed to report prevalence rates and refused or were unable to provide relevant estimates through personal communication. This could result in publication bias, but the exact impact of it is not ascertainable using any exiting methods.

Selection bias

Selection bias can be defined as a distortion in the estimate that results from procedures in which subjects are selected for the study (Porta et al., 2008). In cases of systematic reviews, it can be introduced by omitting studies that differ systematically from included studies. Eligibility criteria, based on judgement over adequacy of specific definitions, were considered arbitrary and have potential for selection bias. The risk of selection bias was reduced by using broad inclusion

criteria and a comprehensive search. In addition, all eligible prevalence estimates were included and further summarised through either narrative synthesis or meta-analyses. To reduce the risk of selection bias, study quality served as an indicator of sources of bias, rather than an inclusion criterion. When possible, study quality was included in meta-regression analyses. As a result of these decisions the task of managing potential information bias posed substantial challenges.

Information bias

Information bias is considered to result from systematic differences between various study groups in the quality of data obtained on exposure, covariates or outcome or flow in estimates resulting from measurement error (Porta et al., 2008). Although the reviewed studies were of variable quality, quality assessment suggested that most studies were of fair or good quality. The most common sources of information bias included limited sample size, heterogeneous case definitions for OA and the use of a variety of definitions and measures of anxiety and depression. Attention to these features in future studies would potentially reduce bias and assist the synthesis and pooling of data – a point that has been made before in the context of systematic reviews of Anderson et al. (2001), Grigsby et al. (2002), Barnard et al. (2006) and Mitchell et al. (2011).

Information bias could have impacted on the findings presented in this review. As suggested by the methodological quality assessment, the issue of sample size calculations could impact on included estimates of specific anxiety disorders. The Egger's test supported large effects for GAD and PTSD prevalence estimates in smaller studies indicating prevalence of those anxiety disorders is likely to be overestimated in small samples. Similar observations emerged from

previous systematic reviews (e.g. Anderson et al., 2001, Mitchell et al., 2011).

Different studies used varied definitions of OA/joint pain. This made comparisons across studies challenging. The clarity over the impact of this problem on pooled estimates could be clarified by comparing prevalence rates across groups with different severity of symptoms or anatomical site of pain. Unfortunately, as included studies were mostly cross-sectional, with limited availability of additional data, this was not possible.

As with defining OA, there are inconsistencies in the definitions of depression and anxiety used by individual studies, which pose challenges when comparing the identified prevalence estimates. This problem has been previously recognised by Regier et al. (1998) in their epidemiological study on the prevalence of mental disorders, where the authors highlight the problem of heterogeneity of prevalence rates and decisions around which symptoms should be of interest. The only way to reconcile this issue of disparate definitions, involves arbitrary a priori decisions. This, however, generates the risk of selection bias.

Whilst self-report questionnaires are based on similar reference standards (Williams et al., 2002) and use similar classifications for symptom severity, prevalence rates of anxiety and depression in people with OA vary across the different tools used. The process of grouping prevalence rates exposes the issue of a lack of consensus over applicability of specific tools in patients with osteoarthritis and compatibly of anxiety and depressive measures used in patients with OA. Consequently, in line with Grigsby et al. (2002), management of the variety of assessment tools and cut-off points was particularly challenging in this study. Sub-group analyses indicated the impact of methods of ascertainment of estimates, but the heterogeneity within sub-group estimates indicated that heterogeneity is multi-

factorial and could not be simply resolved by arbitrary inclusion to specific methods of ascertainment.

Two general observations should be acknowledged regarding sources of anxiety questionnaire data. Anxiety in general was unlikely to be investigated on its own. Consequently, given that HADS includes two short depression and anxiety subscales, it seemed to be a practical choice of tool to use. Another issue is that studies reporting estimates of anxiety were derived from closely related sample frames, namely older people with OA/joint pain recruited from one area of England (North Staffordshire). Relatively low heterogeneity of anxiety estimates could be related to these two factors, but the exact impact of it on the generalisability of these findings is unclear.

2.7.3 Comparison with other populations

Adults with osteoarthritis vs. osteoarthritis-free adults

It is difficult to interpret differences between pooled anxiety and depression prevalence estimates in adults with OA and the general population, due to confounding impacts of age, sex and presence of other co-morbidities. There is evidence from primary general population surveys which suggest that adults with OA might be at an increased risk of depression and anxiety compared to those without OA. A nested case-control study of adults aged 40-79 has shown that 300 people with knee OA had significantly higher mean HADS depression and anxiety scores than 300 age and sex matched controls without OA (O'Reilly et al., 1998). The study has also shown significantly higher frequency of mild to severe depressive and anxiety symptoms (HADS score ≥8) in those with knee OA than their counterparts without OA (O'Reilly et al., 1998). A sample of community-dwelling adults

with joint disorders (including arthritis) compared with a pain free group; revealed that OA may be associated with higher odds of depressive disorders, panic attacks and GAD [McWilliams et al., 2004].

Adults with osteoarthritis vs. with other musculoskeletal diseases

To the author's best knowledge, there are no published systematic reviews and meta-analyses of prevalence rates of depression and anxiety coexisting with other musculoskeletal disease. Evidence from primary studies suggests that prevalence rates of depressive disorders, panic disorder and GAD reported by adults with joint disorders [McWilliams et al., 2004, Munce & Stewart, 2007] are 1% to 3% lower than these coexisting with back problems. In Munce and Stewart's study (2007) women with fibromyalgia reported a prevalence of depressive disorders of 23.7%, considerably higher than 11.0% prevalence of depressive disorders in women with arthritis/rheumatism. The same pattern, but less considerable was found for men. The discrepancy could be due to pain characteristics of these two conditions, but could also be affected by the lack of adjustment for age or a low prevalence (1%) of fibromyalgia in the investigated population.

Adults with osteoarthritis vs. with cardiovascular problems and diabetes (i.e. long-term conditions for which recognition is already advocated)

Two reviews of prevalence rates of depression coexisting with chronic heart failure could be identified [Rutledge et al., 2006, Yohannes et al., 2010]. Yohannes et al. (2010) also identified prevalence rates of anxiety coexisting with chronic heart failure. Rutledge et al. (2006) conducted a non-systematic review but descriptions of included studies were lacking, so results are not directly comparable to this study.

Pooled prevalence rates of depression assessed using higher and lower cut-off points were 14.0% and 38.0% (Rutledge, 2006). Yohannes et al. (2010) also used a non-systematic approach, failed to state eligibility criteria and only partially described the studies included. Prevalence rates of anxiety symptoms ranged from 11.0% to 45.0% and 18.0% of clinically diagnosed anxiety disorders (Yohannes et al., 2010).

Three reviews of depression prevalence in Type 1 diabetes (Barnard et al., 2006), Type 2 diabetes (Ali et al., 2006) and both types combined (Anderson et al., 2001) were identified. A systematic review of prevalence rates of anxiety in both types of diabetes is also available (Grigsby et al., 2002). In reviews by Ali et al. (2006) and Barnard et al. (2006) questionnaires and clinical interview data were aggregated. Mixed settings study designs were included. As sub-group analyses were not provided in the above, a systematic review by Anderson et al. (2001), which did include sub-group analyses, is arguably more useful. The authors reported 11.4% prevalence of depressive disorders assessed with diagnostic interview data (Anderson et al., 2001). This is comparable to the estimated prevalence rates of major depression and dysthymia in this thesis. Anderson et al. (2001) reported 31.0% pooled prevalence rate of depression symptoms, which is 8% higher than the estimate reported in this thesis (Anderson et al., 2001). Pooled prevalence estimates reported for community samples (20%) and mixed clinical settings (32%) suggest that the discrepancy between prevalence rates of depression symptoms could be due to the impact of spectrum bias (Anderson et al., 2001). Spectrum bias occurs when cases within a limited range of a disease spectrum are included, such as medical inpatients or participants from randomised controlled trials (Willis, 2008).

Comparison with the Grigsby et al.'s (2002) review is more challenging as the authors provided incomplete descriptions of included studies and did not perform meta-analyses and sub-group analyses by study settings. Grigsby et al. (2002) reported a pooled prevalence rate of panic disorder (1.3%) coexisting with diabetes that was lower than the estimate presented earlier in this chapter for patients with OA/joint; and a pooled prevalence rate of PTSD (1.2%) was comparable to that estimated for adults with OA/ joint pain (Grigsby et al., 2002). Pooled prevalence of GAD (13.5%) and social phobia (7.3%) were higher than those estimated in people with OA/ joint pain (Grigsby et al., 2002). Notably, 30% of pooled estimates of anxiety disorders were derived from studies with less than 100 participants and 30% from studies with between 100 to 200 participants. A pooled prevalence of elevated anxiety symptoms coexisting with diabetes (39.6%) (Grigsby et al., 2002) was lower than that estimated for adults with OA. This may be due to the inclusion of various questionnaires and cut-off points in the current review. In general, comparing the current study with the reviews for patients with diabetes was challenging, due to discrepancies in the ways the meta-analyses were carried out. Prevalence rates of depressive disorders coexisting with OA and diabetes are likely to be comparable. Seemingly, once the impact of disease spectrum biases and used questionnaires are taken into account, prevalence rates of elevated depression symptoms coexisting with diabetes and OA may well be comparable. Prevalence rates of specific anxiety disorders coexisting with diabetes are likely to be distorted by sample size, and thus, should be interpreted with caution.

2.7.4 Strengths and limitations

Strengths in the current study were: the use of a systematic approach to the

identification and selection of studies; additional search of related articles and attempts to contact authors for missing information; independent data extraction and quality assessment; the use of a clear approach to standardised weighting for pooled prevalence estimates; and rigorous efforts to understand the sources of heterogeneity. On this latter point, comparing two suggested methods of exploring the impact of individual studies, demonstrated that both methods were of limited use, in this particular study at least. An acceptable heterogeneity level can be reached by omitting individual studies, but this method tends to selectively omit large studies. As previously noted (Song et al., 2001), the author of this thesis concludes that the conducted meta-regression analyses of prevalence rates require caution on interpretation. This is particularly due to limited variability or partial unavailability of data for studies entered in meta-regression analyses.

Heavy reliance of clinical interview data on He et al.'s (2008) study should be acknowledged. The current review contributed to this study by providing valuable pooled estimates. Other potential criticism of this work which might arise pertains to the decisions around the selection and organisation of the data. Possible difficulties related to the decisions discussed in section 2.7.2 on page 71 were dealt with by: reporting results with full transparency and highlighting limitations in interpretations, performing sensitivity analyses for the impact of broadly defined 'arthritis', taking a random effect approach, which assumes natural variability between estimates, providing sub-group estimates and quantifying effects of individual factors on heterogeneity.

2.7.5 Implications for clinical practice

Studying the frequency of a health phenomenon cannot be expected to

result in immediate changes to clinical practice without strong evidence from other forms of research (e.g. treatment effectiveness or diagnostic test accuracy). Nevertheless, the descriptive findings in this chapter arguably focus attention towards two important clinical implications. The higher frequency of anxiety symptoms over depression does not appear to be reflected in current management guidance where the emphasis has firmly been on targeted depression case identification and management in patients with physical conditions. The development and routine use of brief instruments for anxiety assessment such as the HADS-A (Zigmond & Snaith, 1983) or the GAD-7 (Spitzer et al., 2006) and ultra-brief instruments such as the GAD-2 (Kroenke et al., 2007) may redress this imbalance.

On the basis of frequency alone, the case for recognition of coexisting anxiety and/or depression appears as strong in the case of OA as for other physical conditions specifically identified in the existing guidelines (e.g. diabetes and cardiovascular disease). The high prevalence of anxiety and depression symptoms together with the high consultation prevalence of OA suggests that a degree of targeting may be needed. One approach is to focus efforts on the recognition of rarer anxiety and depressive disorders. Evidence suggests this approach is useful for treatment decision making as all identified cases would arguably qualify for intervention - the assumption that cannot be made based on questionnaire assessment (Van Rijswijk et al., 2009). On the other hand, as mentioned in the context of recognition of schizophrenia, this approach can be argued to mitigate secondary prevention (Van Os & Delespaul, 2005). It may also require using clinical interview schedules arguably unfeasible for use in primary care (Hanel et al., 2009). Another approach is to distinguish between relatively transient states of

anxiety and depression and more persistent symptoms and investigate the possibility of identifying in the latter cases that may benefit from an intervention.

2.8 CONCLUSIONS

This review provides estimates of the frequency with which anxiety and depression symptoms and disorders occur in adults with OA and joint pain in the community. The 'best estimates' are: major depression 7%; dysthymia 3%; 'mild or worse' depression symptoms 24% and 'moderate or worse' depression symptoms 15%; generalised anxiety disorder 3%; social phobia 3%; panic disorder 3%; panic with agoraphobia 2%; post-traumatic stress disorder 2%; 'mild or worse' anxiety symptoms 45% and 'moderate or worse' anxiety symptoms 21%. The key sources of bias in the available evidence were missing data, the impact of sample size, and the lack of consistency in case definitions and measurement instruments. This might reflect the complexity of the conditions under investigation and the lack of consensus in this field. These are potential areas for improvement in future studies.

Putting to one side the fact that there may be considerable heterogeneity in pooled estimates, this study draws attention to an important implication. Namely, in seeking to identify cases with depressive or anxiety disorders it can be anticipated that there will be as many as 4-15 times that number of individuals who report anxiety or depression symptoms of whom half will be reporting moderate or worse symptoms. The significance of these more common symptoms is clearly of interest for primary care management. In chapter five the extent to which these are likely to be transient or persistent states among a sample of older patients presenting with musculoskeletal pain to primary care will be investigated.

Chapter three: Self-report measures for depression and anxiety symptom screening and assessment in osteoarthritis: a narrative review of selected measurement properties

3.1 INTRODUCTION

The previous chapters have outlined the context for the recognition, measurement and classification of anxiety and depression symptom severity in general practice, and provided estimates of the frequency of these in adults with OA/joint pain. While the reference standards for diagnosis of anxiety and depressive disorders are generally well-established, greater choice and uncertainty exist around which of the many patient-reported self-complete instruments used to ascertain and quantify the severity of anxiety and depression symptoms are most suitable for use in this subpopulation. This chapter identifies the potentially suitable measures in this subpopulation and critically considers the evidence for their reliability, responsiveness, validity, and acceptability to patients and health care professionals.

3.2 RATIONALE OF THE STUDY

Diagnostic tests – those that help discriminate between patients with and without a disease or health condition of interest - are a critical part of health care (Leeflang et al., 2008). Health professionals may be interested in knowing which of the available depression and anxiety questionnaires should be used, how they should be interpreted, and whether the use of those tests has been demonstrated to improve patient outcomes (van Rijswijk et al., 2009). As discussed in chapter one, clinical guidelines suggest that in aiding the assessment of depressive and anxiety

symptoms, self-report instruments serve several purposes covering the recognition of symptoms and the measurement and classification of symptom severity. Their primary purpose is not the diagnosis of an anxiety and depressive disorder *per se* although they may contribute to this process, for example as a basis for initial case identification. The evaluation of self-report instruments for anxiety and depression symptoms must therefore extend beyond diagnostic accuracy to incorporate an evaluation of other properties, including aspects of reliability, validity and interpretability as well as practical considerations of acceptability and feasibility (Nunnally, 1978, Myers & Winters, 2002, Hunsley & Mash, 2010, Fava et al., 2012).

It is known that some of these measurement properties vary across settings and the population in which they are assessed (Bot et al., 2004). The challenge of the current work was to consider evidence for general primary care adults before attempting to critically relate this to subpopulations of older adults and people with musculoskeletal problems, such as osteoarthritis. There are several reasons for believing that the variation in measurement properties noted by Bot et al. (2004) applies here. Firstly, there may be concern over spectrum bias (Furukawa & Guyatt, 2006, Willis, 2008) in so far as evidence of measurement properties derived from tertiary care patients may be particular to a selective part of the spectrum of symptoms (the more severe end or symptoms that are atypical). The same measures may perform differently in primary care patients, where a wider spectrum of symptoms can be expected. Secondly, the recognition and assessment of anxiety and depression symptoms in older age groups, where osteoarthritis becomes increasingly prevalent, is known to present particular challenges. For example, under-recognition of depression in older community-dwelling adults appears to be more prominent than in younger adults (Bowers et al., 1990, Iliffe et al., 1991). In addition,

diagnostic rates of major depression differ across age groups [Volkers et al., 2004] and specific psychological tests might perform differently across younger and older adults [Jeste et al., 2005]. Thirdly, recognition of depressive disorders in primary care patients with musculoskeletal problems might be difficult because of overlap in the somatic aspects of depression and symptoms of musculoskeletal problems (e.g. sleep problems or psychomotor retardation) [Turner & Romano, 1984, Novy et al, 1995, Rosemann et al., 2006]. Indeed, as shown in patients with musculoskeletal pain, using the DSM-IV criteria inclusive of somatic symptoms - regardless of presumed cause increased the prevalence rate of major depression by nearly half [Wilson et al., 2001]. Since somatic-vegetative symptoms are core to anxiety [Blanchard & Blanchard, 1989, Spielberger & Rickman, 1991, Seligman et al., 2001, Sarason & Sarason, 2004, Beesdo et al., 2009], the problem of symptoms overlapping with musculoskeletal disease holds true for anxiety.

In the previous chapter a total of 15 depression and 8 anxiety symptom measures had been used in population and primary care studies of the prevalence of coexisting depression and/or anxiety in adults with OA or joint pain. Table 3.1 overleaf lists these together with the papers describing their use.

Table 3.1 Depression and anxiety symptom measures used in observational studies in OA/ joint pain.

Depression symptom measures (a study in which was used)	Anxiety symptom measures (a study in which was used)
AIMS-D (Creamer, 1999, Dexter, 1994)	**AIMS-A** (Creamer, 1999)
BDI-13 (Viinamäki, 2002)	
BDI-21 (Williams, 2004)	
CESD- 7 (Tsai, 2005)	
CESD- 8 (Polsky, 2005)	
CESD-10 (Szoeke, 2008)	
CESD-20 (Barberger-Gateau,1992, Brandt, 2000, Scudds, 2000, Wilcox, 2000, Kramer, 2002, Baker, 2003, Nour, 2005, Ferreira, 2006, Kalichman, 2007, Maly, 2007, Allen, 2008, Sale, 2008)	
4 DSQ-D (Reilingh, 2008)	**4DSQ-A** (Reilingh, 2008)
GADS-D (Menz, 2006)	**GADS-A** (Menz, 2006)
GDS-15 (Woo, 1994, Niti, 2007, Appelt, 2007, Schram, 2008)	
GDS- 30 (Martin, 1996, Leveille, 2007)	
HADS-D (O'Reilly, 1998, Badcock, 2002, Peat, 2006a, Hill, 2007, Spies-Dorgelo, 2007, Wood, 2007, Gignac, 2008, Mallen, 2008)	**HADS-A** (O'Reilly, 1998, Memel, 2000, Badcock, 2002, Peat, 2006a, Hill, 2007, Spies-Dorgelo, 2007, Mallen, 2008)
IRGL-6 (Hopman-Rock, 1997)	**IRGL-A** (Hopman-Rock, 1997)
Negative affect scale^ (Fisher, 2004)	
PHQ-2 (Figaro, 2005, Mallen, 2008)	
PHQ-9 (Rosemann, 2007)	
	POMS (Hampson, 1996)
	STAI-20 (Williams, 2004)
	STAI-40 (Maly, 2007)

Note: AIMS-A/D - Arthritis Impact Measurement Scales-anxiety/depression subscale; BDI - Beck Depression Inventory; CESD - Center for Epidemiologic Studies Depression Scale; 4DSQ-A/D - Four dimensional Symptoms Questionnaire-anxiety/depression subscale; GADS-A/D - Goldberg Anxiety and Depression Scale-anxiety/depression subscale; GDS -Geriatric Depression Scale; HADS-A/D - Hospital Anxiety and Depression Scale- /depression subscale; IRGL-A/D - Influence of Rheumatic Diseases on Health and Lifestyle-anxiety/depression subscale; PHQ - Patient Health Questionnaire; POMS - Profile of Mood States; STAI - State-Trait Anxiety Inventory;
^ - the 7-item negative affect scale formulated from the CES-D.

A comprehensive evaluation of the measurement properties, characteristics and performance within the target population of interest (older adults with osteoarthritis) for all the measures is beyond the scope of this thesis. Instead, this chapter will focus on comparing the measurement properties and characteristics of a selection of these and their suitability for use in older adults with joint pain.

The three depression measures selected were the Beck Depression Inventory (BDI-II: Beck et al., 1996, BDI-PC: Beck & Steer, 1997), the Hospital

Anxiety and Depression Scale – Depression subscale (HADS-D: Zigmond & Snaith, 1983) and the Patient Health Questionnaire (PHQ-2, PHQ-9: Kroenke et al., 2001, Kroenke et al., 2003). The BDI-II, HADS and PHQ-9 are the measures currently recommended for use in UK primary care settings for assessment of severity of symptoms in general adults and in patients with other chronic physical health problems, i.e. coronary heart disease (BMA, GPC, 2009). The two-item depression screener, derived from the PHQ-9 (i.e. PHQ-2: Kroenke et al., 2003), was also recommended for screening in primary care patients at high-risk of depression (NICE, 2009a, 2009b). BDI-II, PHQ-9 and HADS together with CESD and GDS, were regarded by Smarr and Keefer (p. 454, 2011) as the "*most relevant* [measures] *for the assessment of depression in the context of rheumatology clinical and/or research practice*" although only the BDI-II has been recommended for outcome evaluation in chronic pain clinical trials (Dworkin et al., 2005). All four tools are available as clinical Read terms (i.e. Patient health questionnaire score (388f.), depression screening using questions (6896.), HAD scale: depression score (388P.), Beck depression inventory second edition score (388g.)), when searched with reference to the NHS Information Authority ((NHSIA) 2000) devised Clinical Terms Version 3. The BDI for Primary Care (BDI-PC), now known as Fast Screener (BDI-FS), was included for comparison with the PHQ-2.

For anxiety symptom assessment, the selection is more difficult. There is a lack of agreed authoritative recommendations, reflecting the more limited research evidence in this field. Detailed recommendations are unavailable from either the Quality and Outcomes Framework contract (BMA, GPC, 2009) or NICE guidelines on generalised anxiety disorders, and panic with and without agoraphobia (NICE, 2007). Other NICE anxiety guidelines deal only with narrow disorders such as screening

for Post-Traumatic Stress Disorder (National Collaborating Centre for Mental health (NCC-MH), 2005) and Obsessive-Compulsive Disorder (NICE, 2005). A variety of approaches has been suggested in published research articles and opinion pieces. For screening and case identification these include using an anxiety subscale on the Symptoms Checklist-90-R combined with the SF-36 (Gilbody et al., 2001), or using broad screening tools (e.g. Primary Care Evaluation for Mental Disorders) and brief measures (e.g. General Health Questionnaire) to assess general distress and an anxiety subscale on the Brief Symptom Inventory-18 (Lang & Stein, 2002). For measuring anxiety in primary care they include using anxiety subscales (e.g. HADS-A) and general distress scales (e.g. Kessler Psychological Distress Scale, Somatic and Psychological Health Report) (Hickie et al., 2002).

Recently, the HADS-A (Zigmond & Snaith, 1983) and the Generalised Anxiety Disorder Scale (GAD-7) (Spitzer et al., 2006) have been both reported as useful for assessment of anxiety disorders in primary care medically ill patients (Levenson et al., 2010, Roy-Byrne et al. 2009) although only the HADS-A was included in the recent review of anxiety measures for rheumatologic populations by Julian (2011). The GAD-7 has been used in NHS mental health programmes such as Improving Access to Psychological Therapies (IAPT) (Department of Health, 2008 in Ross, 2010), which included 100 UK Primary Care Trusts (Ross, 2010). IAPT was launched in year 2008 (Lester & Glasby, 2010) following Lord Layard's report, showing poor access to psychological therapies (Centre for Economic Performance's Mental Health Policy Group, 2006). The apparent receptiveness to the recently developed GAD-7 (Spitzer et al., 2006) might be related to the fact that the GAD scale was derived from the PRIME-MD. Both the HADS-A and GAD-7 are available through NHS Read terms (NHSIA, 2000): Generalised anxiety disorder 7 item score (Read code: XaA5Z) and HAD scale: anxiety scale (Read

code: 388N). The GAD-2 (Kroenke et al., 2007) offers promising potential for quick, targeted case-identification of primary care patients at risk of possible/probable anxiety problems. It was decided that the current chapter would therefore focus on these two measures - the Hospital Anxiety and Depression Scale-Anxiety subscale (HADS-A) and the Generalised Anxiety Disorder tool (GAD-2, -7).

3.3 AIM AND OBJECTIVES

The **overall aim** of this chapter was to understand the comparative strengths and weaknesses of recommended patient-reported, self-complete, condition-specific measures for use in assessing anxiety and depression symptoms in patients presenting to primary care with osteoarthritis.

In the context of the thesis, this would also provide the opportunity to evaluate the absolute and relative performance of the HADS – the measure used in empirical analyses presented in subsequent chapters of this thesis.

Specific objective is:

- To conduct a narrative synthesis of evidence on key measurement properties and characteristics and suitability of selected depression (BDI (BDI-II and BDI-PC versions), HADS-D, PHQ (2- and 9-item versions)) and anxiety (GAD (2- and 7-item versions), HADS-A) symptom measures

3.4 METHODS

3.4.1 Evaluation framework

In this study, the framework for evaluating self-report instruments for anxiety and depression symptoms was constructed from traditional psychometric

(e.g. Streiner, 2003) and recent clinimetric approaches (Mokkink et al., 2010). Four main clinimetric aspects of health related measures, agreed through a consensus of 43 clinimetric experts, are: reliability (internal consistency, reliability and measurement error), validity (content validity, criterion validity and construct validity), responsiveness and interpretability (Mokkink et al., 2010). All of these properties are broadly comparable to psychometrics and some of them were considered in the current work, with the addition of acceptability. Acceptability was added as limited acceptability on assessment measures might have practical implications, such as avoidance of its use (Dowrick et al., 2009). Selection of indicative evidence was based on previous literature, with the choice of a method of quantifying concurrent validity being most challenging (see Table C.1.2 in Appendix C.1 on page 355 for a summary of possible methods). As a means of managing the variety of methods to quantify concurrent validity, evidence of rule-out and rule-in accuracy was desired, and thus, the review focused on likelihood ratios (Deeks & Altman, 2004) as they appear to be more commonly used and are considered to be prevalence independent (Deeks & Altman, 2004). Each of the measurement properties and characteristics covered in the used framework is briefly described in Table 3.2 overleaf.

Table 3.2 Adapted evaluation framework.

Dimension	Definition	Components	Definition	Indicative types of evidence
Feasibility	The quality of being practical and possible	Administration burden (Bot et al., 2004)	-	Ease of the scoring method (adequate: symptoms count/ severity or simple algorithms, inadequate: complex formulas)
		Interpretability (Bot et al., 2004)		Interpretation of scores (adequate: underlying reasons for system classification and its relationship with classification defined by a gold standard)
		Accessibility (Julian, 2011)	-	Financial burden (adequate: free, inadequate: charges)
		Readability and comprehension (Bot et al., 2004)	-	Required literacy level (adequate: easy or average, inadequate: difficult)
		Time to administer (Bot et al., 2004)	-	Time taken to complete (adequate: less than 10 minutes)
Reliability	The extent to which the tool is free of measurement error, i.e. the difference between a measured value and its true value (Mokkink et al. 2010)	Internal consistency	The extent to which items in a(sub) scale are interrelated, thus measuring the same construct (Terwee et al., p. 39, 2007)	• **Cronbach's alpha:** alphas 0.70 to 0.80 adequate for research and ≥ 0.90 adequate for clinical purposes (Bland & Altman, 1997) • **Item-total correlation**: correlations in the approximate range of 0.30–0.70 regarded adequate (Streiner & Norman, 2008) • **Factor analyses**: exploratory or confirmatory factor analysis tested (Porta et al., p. 92, 2008, Kline, 1994), adequate if consistent across populations and no issues with item loading emerged
		Test-retest- reliability	The measure's stability over time (Bot et al., 2004), whilst no changes in depression or anxiety occurred	• **Intra-class correlation**: ICC > 0.70 regarded adequate, if a time interval and 95% CI are reported (Boot et al., 2004) • **Correlations between presentations:** correlations r ≥ 0.80 repeated over one to two weeks and r ≥ 0.70 over one month are taken to represent evidence of adequate reliability (Meades & Ayers, 2010)

Table 3.2 cont. Adapted evaluation framework.

Dimension	Definition	Components	Definition	Indicative types of evidence
Construct validity	Construct validity concerns the degree of the measurement's compatibility with theoretical concepts describing the construct under study (Porta et al., 2008)	Convergent validity	The extent to which a scale correlates with other measures that assess the same construct (Porta et al., 2008)	Correlations with measures to assess the same construct. Cohen's criteria used to evaluate effect size of r (r=0.1, small, r=0.3, medium, r=0.5, large (Meades & Ayers, 2010), with a large effect being adequate
		Compatibility with functional status (Kroenke et al., 2001)	Interpretability of the results of assessment in a general context of well-being	Any correlations with general functioning measures considered to be adequate
		'Somatic bias'†	A systematic misclassification of pain and somatic symptoms of depression/anxiety	Degree of somatic symptom variance on the measure, which is not necessarily related to depression/anxiety, but reflects somatic disease. Adequacy: no clear criteria available
Criterion validity	The degree of correlation between the results of an external criterion and the tool (Porta et al., 2008)	Concurrent validity	The correlation between the measure of interest and an external reference standard at the same point in time (Porta et al., 2008)	**Diagnostic accuracy** expressed by the power of changing pre-test probability into the post-test probability (Ahrens & Pigeot, 2005): positive likelihood ratio (LR+) and negative likelihood ratio (LR-), with LR+ above 10 and LR- below 0.10 indicating a large change, i.e. strong evidence for adequate rule-in and rule-out accuracy respectively (Jaeschke et al., 1994, Stengel et al., 2003, Deeks & Altman, 2004)
Responsiveness to change	An ability to capture changes in scores following intervention aiming for symptom reduction (Hays & Hadorn, 1992)	-	-	Distribution-based approaches (i.e. Effect size (ES) standardised response mean (SRM), responsiveness index (RI)) and anchor-based approaches (i.e. use of an external clinical or patient-based criterion to assign subjects into groups reflecting the scope of changes) (Revicki et al., 2008) Cohen 'rule of thumb' used for interpreting ES of change (<0.20: trivial, ≥0.20 and <0.50: small, ≥0.50 to <0.80: moderate, ≥0.80: large) (Revicki et al., 2008), with any change regarded adequate
Acceptability†	Adequacy of the practice of assessment to satisfy a need for use	To patients	-	Quantitative/qualitative analyses of patients' perception Adequacy: no clear criteria available
		To clinicians	-	Quantitative/qualitative analyses of clinicians' perception Adequacy: no clear criteria available

Note: †- given a lack of consensus on interpretation, adequacy could only be judged subjectively.

3.4.2 Literature search and narrative synthesis

3.4.2.1 Search strategy

Published English-language studies reporting primary data (i.e. primary studies or reviews) on one or more measurement property or characteristic were sought through searching 7 electronic databases - Cochrane library, PsycInfo, CINAHL, MEDLINE, EMBASE, CSA illumina and Web of Science - from inception to November 2011 using the following key groups of terms:

1. data synthesis/ recommendations (e.g. review* OR recommend*)
2. depression/ anxiety measures (e.g. depress* OR anxiet*)
3. psychometric/clinimetric properties (e.g. diagnos* OR validity OR reliability)
4. primary care adults (e.g. primary and care)
5. community-dwelling/primary care elderly (e.g. 'old*' AND 'comm* OR prim*')
6. musculoskeletal complaints (e.g. osteoarthrit* OR musculoskelet*)

(Please refer to Box C.2.1 in Appendix C.2 on page 358 for the search strategy)

In addition, the Centre for Review and Dissemination, Cochrane Library, MEDLINE and EMBASE electronic databases were searched for reviews and expert recommendations, from inception to November 2011. Relevant studies known to the authors of the review were also searched and evidence from these and the studies contained therein was extracted.

3.4.2.2 Data extraction

Data was extracted by one reviewer and included: study author, population, setting, sample size, depression and/or anxiety measure and the findings pertinent to the measurement properties and characteristics under investigation.

3.4.2.3 Data analyses and presentation

All data was narratively synthesised and presented separately for depression measures and anxiety measures with the exception of data on acceptability which is presented in one combined section. Where studies provided the necessary figures, likelihood ratios were calculated in Excel, using the equation provided in Table C.1.2 (Appendix C.1, p. 355), to assist the evaluation of criterion validity. To explore relatively high or low likelihood ratios a possibility of bias was considered (Whiting et al., 2003) and 95% CI confidence intervals were estimated using a Diagnostic Test calculator (Schwarz, 2002). Wider 95% CI were attributed to small sample size and small numbers of observed frequencies in a 2x2 diagnostic accuracy table (Table C.1.1 in Appendix C.1, p. 354), so it was not possible to conclude that any observed difference was not due to chance alone. Using the same system as Ahrens and Pigeot (2005) likelihood ratios were described as large (LR+>10.0, LR-<0.10), moderate (5 -10, 0.1-0.2), small (2-5, 0.2-0.5) or negligible (1-2, 0.5-1.0). For ease of reference, these categories were colour-coded in the tabulated results.

3.5 RESULTS

3.5.1 Identified studies

The BDI-II and -PC, HADS and PHQ-9 and -2 can be found compared in several reviews of measures used in primary care settings, including the UK, and systematic reviews including those focused on the impact of questionnaire assisted screening on depression recognition (Pignone et al., 2002), case identification abilities of depression tools in primary care (Williams et al., 2002) and in older primary care attendees specifically (Watson & Pignone, 2002, Snowden et al., 2009). One study

summarised reviews into the diagnostic utility of depression screening tools (Nease & Malouin, 2003). In total, 42 individual papers containing original data were identified and included for review in this chapter with criterion validity of depression and anxiety measures being the most common focus (see Table 3.3 overleaf). The following sections briefly summarise their findings.

Table 3.3 Included studies of measurement properties of selected depression and anxiety symptom measures.

Measurement property	Depression symptom measures			Anxiety symptom measures	
Acceptability†	Coventry, 2011, Dowrick, 2009, Wood, 2002, Simpson, 2008, Rosemann, 2006			Wood, 2002	
	BDI-II or -PC	**HADS-D**	**PHQ-9 or -2**	**GAD-7 or -2**	**HADS-A**
Feasibility	Beck et al., 1996, Pignone, 2002, Williams, 2002, Nease, 2003, Pearson Education, Inc. 2012	Zigmond, 1983, Williams, 2002, Nease, 2003, GL Education Group, 2012	Pignone, 2002, Williams, 2002, Nease, 2003, Snowden, 2009, Pfizer, Inc., 2012	Williams, 2002, Spitzer, 2006, Garcia-Campayo, 2010, Pfizer, Inc., 2012	Zigmond, 1983, Williams, 2002, Nease, 2003, GL Education Group, 2012
Chronbach's alpha	Beck, 1997, Arnau, 2001, Poole, 2009a	El-Rufaie, 1995, Pallant, 2005, Bunevicius, 2007, Cameron, 2008, Terluin, 2009	Kroenke, 2001, Pinto-Meza, 2005, Cameron, 2008, Han, 2008	Spitzer, 2006, Garcia-Campayo, 2010	Pallant, 2005, Bunevicius, 2007, Terluin, 2009
Item-total correlation	Beck, 1997, Arnau, 2001	Cameron, 2008	Cameron, 2008, Han, 2008, Yeung, 2008	Garcia-Campayo, 2010	-
Factor structure	Arnau, 2001, Harris, 2008, Corbière, 2011	Pallant, 2005, Cameron, 2008	Cameron, 2008	Spitzer, 2006, Garcia-Campayo, 2010	Pallant, 2005
Test-retest reliability	-	Angst, 2008	Lowe, 2004b, Han, 2008, Kroenke, 2001	Garcia-Campayo, 2010	Angst, 2008
Convergent validity	Corbière, 2011	Cameron, 2008	Cameron, 2008, Yeung, 2008	Spitzer, 2006, Garcia-Campayo, 2010	Garcia-Campayo, 2010
Correlation with functional status	Krugh, 1997, Arnau, 2001	-	Kroenke, 2001, Klapow, 2002, Rosemann, 2007	Spitzer, 2006, Kroenke, 2007, Garcia-Campayo, 2010, Ruiz, 2011	-
'Somatic bias'	Harris, 2008, Corbière, 2011	Pincus, 1996	-	-	Pincus, 1996
Criterion validity	Beck, 1997, Poole, 2009b	Wilkinson, 1988, El-Rufaie, 1995, Lam, 1995, Harter, 2001, Watt, 2002, Bunevicius, 2007, Terluin, 2009, Axford, 2010	Whooley, 1997,Spizter,1999, Kroenke, 2001, Arroll, 2003, Kroenke, 2003, Carpacioglu, 2004, Henkel, 2004, Han, 2008,Yeung, 2008, Arroll, 2010, Phelan, 2010, Li 2007	Kroenke, 2007, Garcia-Campayo, 2010	Lam, 1995, El-Rufaie, 1995, Harter, 2001, Watts, 2002, Wetherell, 2007, Bunevicius, 2007, Terluin, 2009, Axford, 2010
Responsiveness	Poole, 2009a	Cameron, 2008, Angst, 2008	Lowe, 2004b, Cameron, 2008	-	-

Note: †- acceptability data considered jointly for any anxiety or depression measures, with the aim to seek an improved understanding of implementation of self-report measures in primary care.

3.5.2 Depression symptom measures

3.5.2.1 Feasibility of BDI-II, HADS-D and PHQ-9

The BDI-II is copyrighted by the author and available for purchase from Pearson Assessment (Pearson Education, Inc., 2012). The HADS is copyrighted and available for purchase from GL assessment (GL Education Group, 2012) and the PHQ-9 is freely available from Pfizer Inc. (2012) (see Appendix C.3, p. 360-363 for copies of the three measures). Table 3.4 overleaf shows characteristics, of the three questionnaires, extracted from Williams et al.'s (2002) and Nease and Malouin's (2003) reviews and checked with the original papers with additional information extracted for the BDI-II (Beck et al., 1996) and the HADS-D (Zigmond & Snaith, 1983).

The BDI-II, PHQ-9 and HADS-D require administration time shorter than consultation time in primary care. All three measures have severity classifications, but underlying reasons for these classifications appear to be either arbitrary or lacking. An advantage of the BDI-II is the easy literacy level required. Advantages of the PHQ-9 and the HADS are their brief forms. The PHQ-9 also allows for symptom counts offering a direct translation into the DSM-IV-TR classification (Nease & Malouin, 2003). Based on this information and superior positive predictive value of the PHQ-9, Nease and Malouin (2003) regarded the PHQ-9 most optimal for primary care patients, but authors of other reviews have refrained from stating the superiority of any tool (Williams, et al., 2002, Pignone et al., 2002, Snowden et al., 2009).

Table 3.4 Feasibility of the BDI-II, HADS-D and PHQ-9.

Instrument (abbreviation)	Number of items (shorter versions)/ scale	Possible score range: severity	Admin. time	Literacy level^	Original time frame of items
Beck Depression Inventory-II (BDI-II)	21 (13,7)/ 0-3	0-63: Minimal (0-13) Mild (14-19) Moderate (20-28) Severe (29-63)	2-5 min	Easy	Past two weeks
Hospital Anxiety and Depression Scale-depression (HADS-D)	7 (-)/ 0-3	0-21: Non-cases (0 - 7) Doubtful† (8 - 10) Definite (≥11)	≤ 2 min	Difficult	Past week
Patient Health Questionnaire-depression (PHQ-9)	9 (8,2)/ 0-3	0-27: None (1-4) Mild (5-9) Moderate (10-14) Moderately severe (15-19) Severe (20-27) Diagnostic algorithm‡	2-5 min	Average	Past two weeks

Note: Admin. - Administration;
^ Classified by Williams et al. (2002) using Fog Formula, into three grade reading levels: 3-5 (Easy), 6-9 (Average), ≥9 (Difficult);
† - score range for doubtful and definite cases, are used as indicative of 'mild' and 'moderate to severe' severity respectively (e.g. Kendrick et al., 2009);
‡ - to establish provisional diagnoses for selected DSM-IV disorders (including symptom count, severity and impact).

3.5.2.2 Reliability of BDI-II and -PC, HADS-D and PHQ-9

The BDI-II exhibited Cronbach's alphas adequate for clinical use, as shown in both patients in primary care (0.94) (Arnau et al., 2001) and with chronic pain (0.91) (Poole et al., 2009a). One of these studies, however, found item-total correlations (0.54 - 0.74) suggesting possible item redundancy and two factors (Arnau et al., 2001), while three factors were found in people with musculoskeletal problems (Corbière et al., 2011, Harris et al., 2008). Overall, internal consistency of the BDI-II is unclear (Table 3.5, p. 99).

A direct comparison of Cronbach's alphas of the HADS-D (0.84) and PHQ-9 (0.83) in adult primary care patients showed similar internal consistency, but

adequate for research use only [Cameron et al., 2008]. This was broadly comparable to results in musculoskeletal complaints (HADS-D: 0.84) and older adult (PHQ-9: 0.86) subpopulations. Further head-to-head comparisons confirmed internal consistency, as shown by adequate item-total correlations of the HADS-D (0.47-0.69) and PHQ-9 (0.42 - 0.65), and established factor structures [Cameron et al., 2008]. Chinese version PHQ-9 showed possible item redundancy (0.52 - 0.85).

Test-retest reliability was rarely studied in the relevant populations (Table 3.5 overleaf), where the PHQ-9 showed adequate test-retest reliability over 48 hours [Kroenke et al., 2001], one week [Lowe et al., 2004b] and three week [Han et al., 2008] intervals. Test-retest reliability of the HADS-D over a four week interval [Angst et al., 2008] is unclear, given only one estimate is available and 95% confidence intervals were not reported.

Table 3.5 Reliability of the BDI-II and -PC, HADS-D and PHQ-9.

	Population/country	Setting	N	Cronbach's alpha	Item-total correlation	Factor structure	ICC (95% CI)	Test-retest correlation coefficient (r)
BDI-II								
Arnau, 2001	Young adult/ U.S.A	Primary care	340	0.94	0.54 - 0.74	2	-	-
Corbière, 2011	Chronic pain/ Canada	Hospital	203	-	-	3	-	-
Harris, 2008	Chronic pain/ Canada	Tertiary care	481	-	-	3	-	-
Poole, 2009a	Chronic pain/ UK	Tertiary care	584	0.91	-	-	-	-
BDI-PC								
Beck, 1997	Adults/ U.S.A	Primary care	56	0.88	0.51 - 0.77	-	-	-
Poole, 2009a	Chronic pain/ UK	Tertiary care	584	0.84	-	-	-	-
HADS-D								
Angst, 2008	MSK/ Switzerland	Tertiary care	273	-	-	-	0.89	-
Bunevicius,2007	Adult/ Lithuania	Primary care	503	0.78	-	-	-	-
Cameron, 2008	Adult/ UK	Primary care	1063	0.84	0.47 - 0.69	1	-	-
El-Rufaie,1995	Adult/ UAE	Primary care	217	0.88	-	-	-	-
Pallant, 2005	MSK/ Australia	Hospital	296	0.84	-	1	-	-
Terluin, 2009	Adult/ Netherlands	Primary care	295	0.83	-	-	-	-
PHQ-9								
Cameron, 2008	Adult/ UK	Primary care	1063	0.83	0.42 - 0.65	1	-	-
Han, 2008	Older adult/ South Korea	Community	1060¥	0.86	0.40	-	-	0.79
Kroenke, 2001	Adult/ U.S.A	Primary care	3000	0.89	-	-	-	0.84π
Löwe, 2004b	Older adult/ U.S.A	Primary care	123	-	-	-	0.81†; 0.96‡	-
Pinto-Meza, 2005	Adult/ Spain	Primary care	375	0.86^	-	-	-	-
Yeung, 2008	Adult/ U.S.A	Primary care	1940	-	0.52 - 0.85	-	-	-

Note: ICC - Intra-class correlation; MSK - with musculoskeletal complaints; UAE - United Arab Emirates;
^ - 0.82 for telephone administered; ¥- of which n=56 used for test-retest reliability; † - the worst-case sample (n=41), i.e. change due to treatment and no control of prior depression; ‡ - the 'best-case' sample (n=82), i.e. the same number of DSM-IV depression symptoms at both assessments; π - administered telephonically.

3.5.2.3 Construct validity of BDI-II and -PC, HADS-D and PHQ-9 and -2

Evidence relating to convergent validity and correlation with functional status is summarised in Table 3.6.

Table 3.6 Construct validity of the BDI-II and -PC, HADS-D and PHQ-9 and -2.

	Population/ country	Setting	N	Convergent validity	Correlation with functional status
BDI-II					
Arnau, 2001	Adult/ U.S.A	Primary care	340	-	Negative correlation with functional status‡
Corbière, 2011	Chronic pain/ Canada	Hospital	206	With CESD (r=0.66-0.72)	-
Han, 2008	Older adult/ South Korea	Community	1060	With PHQ-9 (r=0.77)	-
Krugh, 1997	Adult/ U.S.A	Primary care	77	-	Negative correlation with AIMS physical function score, general health perception and arthritis impact scores, but no correlation with 50-foot walk time and mean grip strength, AIMS pain and social scores
HADS-D					
Axford, 2010	OA/ UK	Tertiary care	54	-	Negative correlation with disability assessed with the Western Ontario and McMaster Universities Osteoarthritis Index Function scale
Cameron, 2008	Adult/ UK	Primary care	1063	With PHQ-9 (r=0.68)	-
PHQ-9					
Cameron, 2008	Adult/ UK	Primary care	1063	With HADS-D (r=0.68)	-
Han, 2008	Older adult/ South Korea	Community	1060	With GDS (r=0.74) With BDI-II (r=0.77) With CESD (r=0.67)	-
Klapow, 2002	Adult/ U.S.A	Primary care	2466	-	Negative correlation with functional status‡
	Older adult/ U.S.A	Primary care	534	-	Negative correlation with functional status‡
Kroenke, 2001	Adult/ U.S.A	Primary care	3000	-	Negative correlation with functional status ‡
Rosemann, 2007	OA/ Germany	Primary care	1021	-	Negative correlation with the AIMS2-SF
PHQ-2					
Kroenke, 2003	Adult/ U.S.A	Primary care	3000	-	Negative correlation with functional status‡
Li, 2007	Older adults/ U.S.A	Community	8205	-	Negative correlation with functional status‡

Note: AIMS (SF) - Arthritis Impact Measurement Scale (Short version); CESD - Center for Epidemiologic Studies Depression Scale; GDS - Geriatric Depression Scale;
‡ - assessed with the Short Form Health Survey (12 or 20 items versions).

Convergent validity

There is evidence for convergent validity of the PHQ-9 and HADS-D, which correlated (r=0.68) in a large sample of adult primary care patients [(Cameron et al., 2008)] (Table 3.6, p. 100). The PHQ-9 and BDI-II were significantly correlated (r=0.77) in a large study of community-dwelling older adults [(Han et al., 2008)]. The BDI-II and Center for Epidemiologic Studies Depression Scale were correlated (r=0.66-0.72) in a small group of hospital patients with musculoskeletal complaints [(Corbière et al., 2011)]. Overall, given consistently large correlations found to across populations, adequate convergent validity of can be assumed.

Correlation with functional status

There is consistent evidence for significant association between increasing PHQ-9 and -2 scores and decreasing functional status of adults, older adults and primary care patients with OA (Table 3.6 on page 100). Significant correlations were also found between BDI-II scores and functional status of primary care adults [(Krugh, 1997, Arnau et al., 2001)]. Likewise, HADS-D scores were significantly associated with functional status of tertiary care patients with OA [(Axford et al., 2010)].

'Somatic bias'

The problem of 'somatic bias' was predominantly investigated for the BDI-II in North American populations [(Callahan et al., 1991, Harris & D-Eon, 2008, Corbière et al., 2011)], with an addition of one UK based study using the HADS-D [(Pincus et al., 1996)]. No relevant data for the PHQ-9 could be identified. A brief summary of the identified studies is displayed in Table 3.7 overleaf.

Table 3.7 Summary of identified studies into somatic bias of BDI-II and HADS-D.

	Population/ country	Setting	N	Selected findings	Relevant conclusions
HADS-D					
Pincus, 1996	RA cases Matched controls/ UK	Cases: Tertiary care Controls: ?	Cases: 163 Controls: 115	• A rheumatologist indicated items 1† and 4‡, as closely related to RA status • Cases scored higher than controls on items 1 to 6	The elevation in depression scores are not an artefact of RA, i.e. the tool is relatively free of criterion contamination
BDI-II					
Corbière, 2011	Chronic pain/ Canada	Tertiary care	206	• 3 factor structure found, including somatic • Mean score of the somatic dimension was the highest of the three • Perceived reasons for experiencing each symptom varied across 3 factors	• High endorsement of somatic items greatly contributed to overall scores • A sub-question on cause of somatic symptoms is needed
Harris, 2008	Chronic pain/ Canada	Tertiary care	481	• 3 factor structure found, including somatic • Item-total correlations for somatic items ranged from 0.34 to 0.58 • Somatic items most likely to be endorsed • All dimensions were equally associated with subjective pain experience	• Inflation of depression score is unlikely • A need for inclusion of somatic items can be assumed

Note: RA - rheumatoid arthritis;

†- item 1: "I enjoy things as much as I used to"; ‡ - item 4: "I feel as if I am slowed down".

Given an apparent lack of consensus on quantifying the degree of 'somatic bias' and its impact on the tool's general utility, the identified evidence is not directly comparable. Overall, the limited number of studies that attempted to ascertain overlap between chronic pain and somatic symptoms of depression,

found some degree of somatic symptom variance on the HADS-D [Pincus et al., 1996] and BDI-II [Harris & D-Eon, 2008, Corbière, 2011], which is not necessarily related to depression, but may reflect somatic disease. Furthermore, two studies found that BDI-II somatic symptom variance would explain a large part of inflation in scores on this scale in patients with musculoskeletal problems [Harris & D-Eon, 2008, Corbière, 2011]. However, the impact of removing BDI-II and HADS-D somatic items, on their ability to discriminate between patients with and without depression has not been investigated, i.e. the extent to which somatic symptoms on self-report measures can be informative of major depression is unclear.

3.5.2.4 Criterion validity of BDI-II and -PC, HADS-D and PHQ-9 and -2

Likelihood ratios for each of the depression measures and in all populations are listed in Table 3.8 (please refer to Table C.4.1 in Appendix C.4 on page 365 for data extracted to estimate likelihood ratios). For ease of comparability across studies the findings have been grouped according to the specific depression measure and version (e.g. the BDI-II) and criterion (e.g. major depression) being investigated and then ordered by cut-point.

Table 3.8 Criterion validity of the BDI-II and -PC, HADS and PHQ-9 and -2.

	Population/ country	Setting	N	Criterion	Reference standard	Cut-off	Likelihood ratios*	
							LR+	LR-
BDI-II								
Poole, 2009b	Chronic Pain/ UK	Tertiary care	36	Major depression + recurrent depressive disorder	SCID (DSM-IV)	15	2.5	0.00
Poole, 2009b	Chronic Pain/ UK	Tertiary care	36	Major depression + recurrent depressive disorder	SCID (DSM-IV)	22O	8.9	0.12
Poole, 2009b	Chronic Pain/ UK	Tertiary care	36	Major depression + recurrent depressive disorder	SCID (DSM-IV)	25	100#	0.27
BDI-PC								
Beck,1997	Adult/ U.S.A	Primary care	56	Major depression	PRIME-MD (DSM-III-R)	6O	17.0#	0.18
HADS								
Bunevicius, 2007	Adult/ Lithuania	Primary care	503	Major depression	MINI (DSM-IV-TR)	6O	2.6	0.29
Harter, 2001	MSK/ Germany	Tertiary care	206	Major depression + recurrent major depression +dysthymia	M-CIDI (DSM-IV)	16R	2.7	0.31
Terluin, 2009	Adult/ Netherlands	Primary care	295	Moderate + severe major depressive disorder	CIDI (DSM-IV)	8R	1.4	0.13
Terluin, 2009	Adult/ Netherlands	Primary care	295	Moderate + severe major depressive disorder	CIDI (DSM-IV)	11O	1.8	0.34
Terluin, 2009	Adult/ Netherlands	Primary care	295	Moderate + severe major depressive disorder	CIDI (DSM-IV)	12O	1.8	0.55
Terluin, 2009	Adult/ Netherlands	Primary care	295	Moderate + severe major depressive disorder	CIDI (DSM-IV)	14	2.7	0.62
Lam, 1995	Older adult/ UK	Primary care	100	Depressive, anxiety, sleep disorders	CIS (DSM-III)	6O	8.7	0.24
Wilkinson,1988	Adult/ UK	Primary care	100	Depressive and anxiety states	SCID (DSM-III)	8RO	6.4	0.12
El-Rufaie,1995	Adult/ UAE	Primary care	217	Depression	CIS (DSM-III)	7O	22.0#	0.35
Axford, 2010	MSK/ UK	Tertiary care	54	Depression	Structured clinical interview (ICD-10)	8R	3.5	0.55
Watts, 2002	Older adult/ UK	Primary care	268	Subclinical mood disorders	GMSA (DSM-III)	8R	2.2	0.85

Table 3.8 cont. Criterion validity of the BDI-II and -PC, HADS and PHQ-9 and -2.

	Population/ country	Setting	N	Criterion	Reference standard	Cut-off	Likelihood ratios*	
							LR+	LR-
PHQ-9								
Phelan, 2010	Older adult/ U.S.A	Primary care	71	Major depression	SCID (DSM-IV)	8	3.6	0.16
Arroll, 2010	Adult/ NZ	Primary care	2642	Major depression	C-CIDI (DSM-IV)	8	5.8	0.16
Phelan, 2010	Older adult/ U.S.A	Primary care	71	Major depression	SCID (DSM-IV)	9	4.4	0.16
Gilbody, 2007	Adults/ UK	Primary care	96	Major depression	SCID (DSM-IV)	9	3.5	0.08
Kroenke, 2001	Adult/ U.S.A	Primary care	580	Major depression	Overview of SCID-I (DSM-IIIR) + PRIME-MD (DSM-IV)	9	5.9	0.06
Phelan, 2010	Older adult/ U.S.A	Primary care	71	Major depression	SCID (DSM-IV)	10R	3.5	0.46
Gilbody, 2007	Adults/ UK	Primary care	96	Major depression	SCID (DSM-IV)	10R	4.2	0.11
Kroenke, 2001	Adult/ U.S.A	Primary care	580	Major depression	Overview of SCID-I (DSM-IIIR) + PRIME-MD (DSM-IV)	10R	7.3	0.14
Arroll, 2010	Adult/ NZ	Primary care	2642	Major depression	C-CIDI (DSM-IV)	10R	8.4	0.28
Yeung, 2008	Adult/ U.S.A	Primary care	184	Major depression	CB-SCIDI-I/P(DSM-IV)	15R	17.0	0.09
Kroenke, 2001	Adult/ U.S.A	Primary care	580	Major depression	Overview of SCID-I (DSM-IIIR) + PRIME-MD (DSM-IV)	15	13.6	0.34
Arroll, 2010	Adult/ NZ	Primary care	2642	Major depression	C-CIDI (DSM-IV)	15R	15.0	0.57
Phelan, 2010	Older adult/ U.S.A	Primary care	71	Minor + major depressive disorder	SCID (DSM-IV)	6	2.5	0.34
Phelan, 2010	Older adult/ U.S.A	Primary care	71	Minor + major depressive disorder	SCID (DSM-IV)	8O	4.4	0.28
Phelan, 2010	Older adult/ U.S.A	Primary care	71	Minor + major depressive disorder	SCID (DSM-IV)	10R	5.1	0.47
Han, 2008	Older adult/ U.S.A	Community	1060	Depressive disorder	MINI (DSM-IV)	2	2.2	0.07
Han, 2008	Older adult/ U.S.A	Community	1060	Depressive disorder	MINI (DSM-IV)	5O	3.6	0.26
Han, 2008	Older adult/ U.S.A	Community	1060	Depressive disorder	MINI (DSM-IV)	10R	7.7	0.49

Table 3.8 cont. Criterion validity of the BDI-II and -PC, HADS and PHQ-9 and -2.

	Population/ country	Setting	N	Criterion	Reference standard	Cut-off	Likelihood ratios*	
							LR+	LR-
PHQ-2 (scale)								
Phelan, 2010	Older adult/ U.S.A	Primary care	71	Major depression	SCID (DSM-IV)	1	2.2	0.21
Kroenke, 2003	Adult/ U.S.A	Primary care	580	Major depression	SCID (DSM-III)	1	2.4	0.04
Arroll, 2010	Adult/ NZ	Primary care	2642	Major depression	C-CIDI (DSM-IV)	1	2.4	0.07
Phelan, 2010	Older adult/ U.S.A	Primary care	71	Major depression	SCID (DSM-IV)	3R	4.2	0.44
Kroenke, 2003	Adult/ U.S.A	Primary care	580	Major depression	SCID (DSM-III)	3	8.3	0.19
Arroll, 2010	Adult/ NZ	Primary care	2642	Major depression	C-CIDI (DSM-IV)	3R	7.6	0.42
Phelan, 2010	Older adult/ U.S.A	Primary care	71	Major depression	SCID (DSM-IV)	4	5.4	0.67
Kroenke, 2003	Adult/ U.S.A	Primary care	580	Major depression	SCID (DSM-III)	4	17.0	0.48
Arroll, 2010	Adult/ NZ	Primary care	2642	Major depression	C-CIDI (DSM-IV)	4	10.0	0.62
Phelan, 2010	Older adult/ U.S.A	Primary care	71	Minor + major depressive disorder	SCID (DSM-IV)	1	2.5	0.27
Phelan, 2010	Older adult/ U.S.A	Primary care	71	Minor + major depressive disorder	SCID (DSM-IV)	3R	5.3	0.52
Phelan, 2010	Older adult/ U.S.A	Primary care	71	Minor + major depressive disorder	SCID (DSM-IV)	4	17.0	0.66
Kroenke, 2003	Adult/ U.S.A	Primary care	580	Any depressive disorder	SCID (DSM-III)	1	2.5	0.14
Kroenke, 2003	Adult/ U.S.A	Primary care	580	Any depressive disorder	SCID (DSM-III)	3	14.0	0.44
PHQ-2 (binary response version)								
Whooley, 1997	Adult/ U.S.A	Primary care	536	Major depression	QDIS (DSM-IV)	'yes' to either R	2.2	0.07
	Age 18-35/ U.S.A	Primary care	51	Major depression	QDIS (DSM-IV)	'yes' to either R	2.4	0.00
	Age 35-64/ U.S.A	Primary care	358	Major depression	QDIS (DSM-IV)	'yes' to either R	2.0	0.10
	Age 65+/ U.S.A	Primary care	127	Major depression	QDIS (DSM-IV)	'yes' to either R	3.2	0.00
Arroll, 2003	Adult/ NZ	Primary care	421	Major depression	C-CIDI	'yes' to either R	2.9	0.05
Arroll, 2003	Adult/ NZ	Primary care	421	Major depression	C-CIDI	'yes' to first	3.0	0.19
Arroll, 2003	Adult/ NZ	Primary care	421	Major depression	C-CIDI	'yes' to second	3.9	0.22
Li, 2007	Older adults/ U.S.A	Community	8205	Minor + major depressive disorder	AUDADIS-IV (DSM-IV)	'yes' to either R	4.4	0.00
Li, 2007	Older adults/ U.S.A	Community	8205	Minor + major depressive disorder	AUDADIS-IV (DSM-IV)	'yes' to first	4.4	0.10
Li, 2007	Older adults/ U.S.A	Community	8205	Minor + major depressive disorder	AUDADIS-IV (DSM-IV)	'yes' to second	5.7	0.18
Li, 2007	Older adults/U.S.A	Community	8205	Minor + major depressive	AUDADIS-IV (DSM-IV)	'yes' to both	5.5	0.27

Table 3.8 cont. Criterion validity of the BDI-II and -PC, HADS and PHQ-9 and -2.

	Population/ country	Setting	N	Criterion	Reference standard	Cut-off	Likelihood ratios*	
							LR+	LR-
PHQ-9 (algorithm)								
Henkel, 2004	Adult/ Germany	Primary care	448	Dysthymia + major depression	CIDI (DSM-IV)	Algorithm	5.5	0.24
Corapcioglu, 2004	Adult/ Turkey	Primary care	1387	Major + minor depressive disorder	Formal interview (DSM-IV)	Algorithm	5.2	0.28
Corapcioglu, 2004	Adult/ Turkey	Primary care	1387	Major depression	Formal interview (DSM-IV)	Algorithm	8.8	0.31
Spitzer, 1999	Adult/ U.S.A	Primary care	585	Major depression	Overview of SCIDI (DSM-IIIR) + PRIME-MD (DSM-IV)	Algorithm	36.0#	0.28

Note: AUDADIS - Alcohol Use Disorder and Associated Disabilities Interview Schedule; CB-SCIDI-I/P - patient version of Chinese-bilingual Clinical Interview schedule for DSM-IVI; (C)CIDI- (computerised) Composite International Diagnostic Interview; CIS - semi-structured clinical interview schedule; DSM - Diagnostic and Statistical Manual of Mental Disorders; GMSA - short Geriatric Mental State Examination; LR- - negative likelihood ratio; LR+ - positive likelihood ratio; MINI - Mini-International Neuropsychiatric Interview; MSK - musculoskeletal complaints; NZ - New Zealand; O - reported as optimal; QDIS- Quick Diagnostic Interview Schedule; PRIME-MD - Primary Care Evaluation of Mental Disorders; R - reported as recommended; SADS - Semi-structured review from the Schedule of Affective Disorders and schizophrenia; SCID - structured clinical interview for DSM disorders; UAE - United Arab Emirates.
- unusually high estimates.
* The colour code used for adopted categories of changes in depressive disorders likelihood was: large - marked in green, moderate - orange, small - red, negligible - white.

The majority of studies (18 out of 26) were conducted for adults in primary care settings, with an additional six studies in older adults and three small studies in people with musculoskeletal complaints. Relative to major depressive disorder (16 studies), depressive disorders categorised for severity (e.g. moderate and severe major depressive disorder) and other types of depressive disorders (e.g. dysthymia), received limited research attention. In a wide range of reference standards used, however, diagnostic criteria used in UK primary care practice (i.e. the ICD-10) were rarely used.

Many studies were small (five used sample sizes smaller than 100). Five studies reported unusually high estimates of LR+ (marked with # in Table 3.8 on page 104). In two of these studies the sample size was smaller than 100 (e.g. Poole et al., 2009b: LR+ 100 (95% CI (1.87, 451), Beck et al., 1997: LR+ 17 (95% CI 4.45, 62.0)). A relatively large estimate was reported by Spitzer et al. (1999), possibly reflecting a degree of incorporation bias, as the index test formed part of the reference test. The exact reason for the unusually high LR+ found by El-Ruifaie and Absood (1995) is unclear. The authors attributed the difference to linguistic factors, but incorporation bias was also possible, as a reference standard was focused on the same constructs as the index test. Overall, increasing cut-off points increased LRs+ at the expense of LRs-, with likelihood ratios typically small or moderate.

Surprisingly few estimates were directly comparable (i.e. studies using the same questionnaire, criterion, reference standard and cut-off) across different populations. Phelan et al.'s (2010) small study in older adults showed consistently lower LR+ estimates on the PHQ-2 scale, than the larger adult primary care studies of Kroenke et al. (2003) and Arroll et al. (2010). In contrast, the nested

analysis in Whoolley et al. (1997) found a little difference in performance of the PHQ-2 (a binary response format) in the oldest participants (65 years or older) and all adults in the sample, and even better performance than in the age group 35-64 years.

There were few directly comparable estimates of diagnostic accuracy across different diagnostic criteria (i.e. based on the same study population, instrument and cut-off, and reference standard). Corapcioglu et al. (2004) found little difference for the PHQ-9 between major depressive disorder and a criterion inclusive of minor depressive disorder. In contrast, in older adults Phelan et al. (2010), found better ruling-out accuracy of PHQ-9 for major depression than for major depressive disorder and a criterion inclusive of minor depressive disorder.

Rule-out (LRs-) and rule-in (LRs+) accuracy apparently differed across specific tools. Overall, the HADS-D and PHQ-2 (a binary response format), were more successful in ruling-out than ruling-in, and the opposite pattern emerged for the PHQ-2 and PHQ-9 scales. The BDI was comparable in both aspects. Two large studies directly compared measures [Whooley et al., 1997, Arrol et al., 2010]. Whooley et al. (1997) have contrasted the PHQ-2 (a binary response format) and BDI-II, with unsatisfactory LRs+ of all the tools and superior rule-out accuracy of the PHQ-2. The study by Arroll et al. (2010) suggests that the PHQ-2 scale is a better at ruling-out than ruling-in and the PHQ-9 is a better in ruling-in than ruling-out major depression in adult primary care patients.

3.5.2.5 Responsiveness of BDI-II and -PC, HADS-D and PHQ-9

There was limited evidence suggesting adequate responsiveness of the PHQ-9 and the HADS-D, with a direct head-to-head comparison suggesting

similar large effect sizes in adult primary care patients (Cameron et al., 2008) (Table 3.9). A study by Poole et al. (2009a) found comparatively small effect sizes for the BDI-II (0.28) and BDI-PC (0.23) in 584 chronic pain patients. Likewise a study by Angst et al. (2008) found a relatively small effect size (0.43) for the HADS-D in 273 chronic pain patients. However, effect sizes appeared to be associated with duration and type of treatment, rather than the population *per se*. For example, increasing effect sizes were associated with longer periods of treatment (Poole et al., 2009a vs. Angst et al., 2008) and the use of specialist mental health treatment, as opposed to psychology informed pain management programs (Cameron et al., 2008 vs. Angst et al., 2008).

Table 3.9 Responsiveness of the BDI-II and -PC, HADS-D and PHQ-9.

	Population/ country	Setting	N	Intervention	Responsiveness
BDI-II					
Poole, 2009a	Chronic pain/ UK	Tertiary care	584	A 16 days multidisciplinary pain management programme	ES=0.28
BDI-PC					
Poole, 2009a	Chronic pain/ UK	Tertiary care	584	A 16 days multidisciplinary pain management programme	ES=0.23
HADS-D					
Angst, 2008	Chronic pain/ Switzerland	Tertiary care	273	A four week in-patient interdisciplinary pain program	ES=0.43, SRM=0.53
Cameron, 2008	Adult/ UK	Primary care	1063	Treatment provided by mental health workers*	ES=1.00
PHQ-9					
Cameron, 2008	Adult/ UK	Primary care	1063	Treatment provided by mental health workers*	ES=0.99
Löwe, 2004b	Older adult/ U.S.A	Primary care	434	Access for up to 12 months to a depression clinical specialist	ES=1.30

Note: ES - effect size; MSK - musculoskeletal complaints; SRM - Standardised Response Mean;
* - a time frame unreported.

3.5.3 Anxiety symptom measures

3.5.3.1 Feasibility of GAD and HADS-A

The GAD-7 is copyright and freely available from Pfizer Inc. (2012). The HADS is copyrighted and available from GL assessment (GL Education Group, 2012) (see Appendix C.3, p. 363-364 for copies of the two anxiety measures). Both tools have a clearly specified classification of severity (Zigmond & Snaith, 1983, Spitzer et al., 2006), but evidence of their compatibility with 'gold standards', across all anxiety disorders seem to be lacking. The required time to administer the GAD-7 (Garcia-Campayo, 2010) and HADS-A (Williams et al., 2002) is compatible with use in a routine GP consultation time slot (Table 3.10). The literacy level required for the HADS-A might pose difficulties for some patients.

Table 3.10 Feasibility of the GAD and HADS-A.

Instrument (abbreviation)	Number of items (shorter versions)/ scale	Possible score range: Severity	Admin. time	Literacy level^	Original time frame of items
Generalised Anxiety Disorder scale (GAD-7)	7 (2)/0-3	0-21: Minimal (0-4) Mild (5-9) Moderate (10-14) Severe (15-21)	2.5 min	Average †	Past 2 weeks
Hospital Anxiety and Depression Scale -anxiety (HADS-A)	7(-)/0-3	0-21: Non-cases (0-7) Doubtful (8-10) Definite (≥11)	≤ 2 minutes	Difficult	Past week

Note: Admin. - Administration;
^ Classified by Williams et al. (2002) using Fog Formula into three grade reading levels: 3-5 (Easy), 6-9 (Average), ≥9 (Difficult);
† - estimated for the original source of items, i.e. the self-reported version of the PRIME-MD (Williams et al, 2002); ‡- score range for doubtful and definite cases, are used as indicative of 'mild' and 'moderate to severe' severity respectively (e.g. Kendrick et al., 2009).

3.5.3.2 Reliability of GAD-7 and HADS-A

The internal consistency of the GAD-7 was found to be adequate for clinical use (Cronbach's α=0.92-0.94) from two studies with a combined sample of 1187 adult primary care patients [Spitzer et al., 2006, Garcia-Campayo et al., 2010] (Table 3.11 overleaf). Although there were no direct head-to-head comparisons of the two measures within the same population the values for HADS-A were somewhat lower (0.78 - 0.83). Similar internal consistency was found for HADS-A in 296 patients with musculoskeletal conditions in hospital setting as observed in adult primary care patients [Pallant & Bailey, 2005].

Limited evidence on other aspects of reliability for either measure in relevant populations was found. In one study, a one-factor structure of the HADS-A was confirmed [Pallant & Bailey, 2005], although one item failed to load adequately. A one-factor structure of the GAD-7 has been consistently supported in two studies with adult primary care patients [Spitzer et al., 2006, Garcia-Campayo et al., 2010], but item-total correlations above 0.68 indicate possible redundancy of some items. One of these studies [Garcia-Campayo et al., 2010] also reported test-retest reliability on the GAD-7 (ICC=0.93) over an interval of one week, but 95% CI were not reported.

Table 3.11 Reliability of the GAD-7 and HADS-A.

	Population/ country	Setting	N	Cronbach's alpha	Item-total correlation	Factor structure	ICC (95% CI)	Test-retest correlation coefficient (r)
GAD-7								
Garcia-Campayo, 2010	Adult/ Spain	Primary care	222	0.94	>0.68	1	0.93	0.84
Spitzer, 2006	Adult/ U.S.A	Primary care	965	0.92	-	1	-	-
HADS-A								
Angst, 2008	Chronic pain/ Switzerland	Tertiary care	273	-	-	-	0.81	-
Bunevicius, 2007	Adult/ Lithuania	Primary care	503	0.78	-	-	-	-
Pallant, 2005	MSK/ Australia	Hospital	296	0.83	-	1 (item 7 loaded poorly)	-	-
Terluin, 2009	Adult/ Netherlands	Primary care	295	0.83	-	-	-	-

Note: ICC - Intra-class correlation; MSK - musculoskeletal complaints.

3.5.3.3 Construct validity of the GAD-7 and HADS-A

Evidence relating to convergent validity and correlation with functional status is summarised in Table 3.12.

Table 3.12 Construct validity of the GAD-7 and HADS-A.

	Population/ country	Setting	N	Convergent validity	Correlation with functional status
GAD-7					
Garcia-Campayo, 2010	Adult/ Spain	Primary care	222	With HADS-A (r=0.90)	Negative correlation with functional status‡
Kroenke, 2007	Adult/ U.S.A	Primary care	965	-	Negative correlation with functional status†
Ruiz, 2011	Adult/ Spain	Primary care	212	-	Negative correlation with functional status‡
Spitzer, 2006	Adult/ U.S.A	Primary care	965	With BDI (r=0.72) With HAM-A (r=0.74)	-
HADS-A					
Garcia-Campayo, 2010	Adult/ Spain	Primary care	222	With GAD-7 (r=0.90)	-

Note: HAM-A - Hamilton Anxiety rating scale;
† - using the 20-Item Short Form Health Survey (SF-20); ‡ - using the World Health Organization Disability Assessment Schedule II (WHODAS II).

Convergent validity

There is evidence for adequate convergent validity of the GAD-7, which was significantly correlated with the Beck Anxiety Inventory (r=0.72) and the Hamilton Anxiety Scale (r=0.74) in a group of 965 adult patients recruited from 15 primary care trusts (Spitzer et al., 2006). The HADS-A and GAD-7 were found to be highly correlated (r=0.90) in a small group of primary care adults (Garcia-Campayo et al., 2010), with a large effect size supporting their convergent validity.

Correlates with functional status

Two studies, with a combined total of 1187 primary care adults, found that increasing GAD-7 scores have been associated with decreasing functioning, as

assessed with the SF-20 (Spitzer et al., 2006, Kroenke et al., 2007) and World Health Organization Disability Assessment Schedule II (Garcia-Campayo et al., 2010). The GAD-7 severity scores have also been shown to predict disability assessed with the WHO-DAS II in 212 primary care adult patients (Ruiz et al., 2011).

'Somatic bias'

No eligible studies were found that presented original evidence on somatic item bias for the GAD-7. 'Somatic bias' was, however, examined for the HADS-A (Pincus et al., 1996). The anxiety scale showed a difference between cases (163 patients with rheumatoid arthritis from an outpatient clinic) and 115 matched controls on items 1 ("I feel tense or wound up") and 4 ("I can relax and sit at ease"), where item 4 was indicated as a possible contaminator by the rheumatologist. Whilst the study authors are inconclusive about this finding, 'somatic bias' seems possible.

3.5.3.4 Criterion validity of GAD-7 and -2 and HADS-A

Likelihood ratios for each of the anxiety measures and in all populations are listed in Table 3.13 overleaf (see Table C.4.2 in Appendix C.4 on page 370 for data extracted to estimate likelihood ratios). The majority of evidence comes from adult primary care patients in the U.S.A, with only two relatively small studies in older adults and two in musculoskeletal disease populations.

Table 3.13 Criterion validity of the GAD-7 and -2 and HADS.

	Population/ Country	Setting	N	Criterion	Reference standard	Cut-off	Likelihood ratios*	
							LR+	LR-
GAD-2								
Kroenke, 2007	Adult/ U.S.A	Primary care	965	GAD	SCID (DSM-IV)	2	2.6	0.08
Kroenke, 2007	Adult/ U.S.A	Primary care	965	GAD	SCID (DSM-IV)	3R	5.1	0.17
Kroenke, 2007	Adult/ U.S.A	Primary care	965	Panic disorder	SCID (DSM-IV)	2	2.5	0.14
Kroenke, 2007	Adult/ U.S.A	Primary care	965	Panic disorder	SCID (DSM-IV)	3R	4.1	0.30
Kroenke, 2007	Adult/ U.S.A	Primary care	965	PTSD	SCID (DSM-IV)	2	2.4	0.22
Kroenke, 2007	Adult/ U.S.A	Primary care	965	PTSD	SCID (DSM-IV)	3R	3.1	0.51
Kroenke, 2007	Adult/ U.S.A	Primary care	965	Social phobia	SCID (DSM-IV)	2	2.3	0.22
Kroenke, 2007	Adult/ U.S.A	Primary care	965	Social phobia	SCID (DSM-IV)	3R	3.6	0.35
Kroenke, 2007	Adult/ U.S.A	Primary care	965	Any anxiety disorder	SCID (DSM-IV)	2	2.9	0.20
Kroenke, 2007	Adult/ U.S.A	Primary care	965	Any anxiety disorder	SCID (DSM-IV)	3R	5.2	0.40
GAD-7								
Kroenke, 2007	Adult/ U.S.A	Primary care	965	GAD	SCID (DSM-IV)	5	2.2	0.05
Garcia-Campayo, 2010	Adult/ Spain	Primary care	212	GAD	Clinically diagnosed (DSM-IV-TR)	8	6.6	0.08
Kroenke, 2007	Adult/ U.S.A	Primary care	965	GAD	SCID (DSM-IV)	10R	4.9	0.13
Garcia-Campayo, 2010	Adult/ Spain	Primary care	212	GAD	Clinically diagnosed (DSM-IV-TR)	10R	13.0#	0.14
Garcia-Campayo, 2010	Adult/ Spain	Primary care	212	GAD	Clinically diagnosed (DSM-IV-TR)	14	99.9	0.38
Kroenke, 2007	Adult/ U.S.A	Primary care	965	Panic disorder	SCID (DSM-IV)	5	2.1	0.11
Kroenke, 2007	Adult/ U.S.A	Primary care	965	Panic disorder	SCID (DSM-IV)	10R	3.9	0.14
Kroenke, 2007	Adult/ U.S.A	Primary care	965	PTSD	SCID (DSM-IV)	5	2.1	0.18
Kroenke, 2007	Adult/ U.S.A	Primary care	965	PTSD	SCID (DSM-IV)	10R	3.5	0.37
Kroenke, 2007	Adult/ U.S.A	Primary care	965	Social phobia	SCID (DSM-IV)	5	2.4	0.22
Kroenke, 2007	Adult/ U.S.A	Primary care	965	Social phobia	SCID (DSM-IV)	10R	3.6	0.35
Kroenke, 2007	Adult/ U.S.A	Primary care	965	Any anxiety disorder	SCID (DSM-IV)	5	2.4	0.16
Kroenke, 2007	Adult/ U.S.A	Primary care	965	Any anxiety disorder	SCID (DSM-IV)	10R	5.5	0.20

Table 3.13 cont. Criterion validity of the GAD-7 and HADS.

	Population/ country	Setting	N	Criterion	Reference standard	Cut-off	Likelihood ratios*	
							LR+	LR-
HADS								
Wetherell, 2007	Older adult/ U.S.A	Primary care	37	GAD	ADIS-IV (DSM-IV)	8R	2.9	0.05
Bunevicius, 2007	Adult/ Lithuania	Primary care	503	GAD	MINI (DSM-IV-TR)	9O	2.8	0.33
Bunevicius, 2007	Adult/ Lithuania	Primary care	503	Panic disorder	MINI (DSM-IV-TR)	11O	4.4	0.00
Bunevicius, 2007	Adult/ Lithuania	Primary care	503	Social phobia	MINI (DSM-IV-TR)	9O	2.6	0.08
Bunevicius, 2007	Adult/ Lithuania	Primary care	503	Social phobia, panic disorder, GAD	MINI (DSM-IV-TR)	9O	3.1	0.31
Terluin, 2009	Adult/ Netherlands	Primary care	295	Panic disorder, agoraphobia, and social phobia	CIDI (DSM-IV)	8R	1.3	0.07
Terluin, 2009	Adult/ Netherlands	Primary care	295	Panic disorder, agoraphobia, and social phobia	CIDI (DSM-IV)	13O	1.9	0.49
Lam, 1995	Adult/ Chinese	Primary care	100	Depressive, anxiety, sleep disorders	CIS (DSM-III)	3	3.9	0.40
Terluin, 2009	Adult/ Netherlands	Primary care	295	Any anxiety disorder	CIDI (DSM-IV)	8R	1.3	0.07
Terluin, 2009	Adult/ Netherlands	Primary care	295	Any anxiety disorder	CIDI (DSM-IV)	13O	2.2	0.50
Harter, 2001	MSK/ Germany	Tertiary care	206	Any anxiety disorder	M-CIDI (DSM-IV)	17	2.7	0.35
El-Rufaie, 1995	Adult/ UAE	Primary care	217	Anxiety	CIS (DSM-III)	8	5.4	0.34
Axford, 2010	MSK/ UK	Tertiary care	54	Anxiety	Structured clinical interview (ICD-10)	8R	4.6	0.15
El-Rufaie, 1995	Adult/ UAE	Primary care	217	Anxiety	CIS (DSM-III)	9O	8.9	0.37
Watts, 2002	Older adult/UK	Primary care	268	Subclinical mood disorders	GMSA (DSM-III)	8R	2.0	0.61

Note: ADIS - Anxiety Disorders Interview Schedule for DSM-IV; CIS - semi-structured clinical interview schedule; CIDI - Composite International Diagnostic Interview; DSM - Diagnostic and Statistical Manual of Mental Disorders; GAD - Generalised Anxiety Disorders; GMSA - short Geriatric Mental State Examination; ICD - International Classification of Diseases; LR- - negative likelihood ratio; LR+ - positive likelihood ratio; M-CIDI - Munich-Composite International Diagnostic Interview; MINI - Mini-International Neuropsychiatric Interview; MSK - musculoskeletal complaints; O - according to authors of a study optimal cut-off point; R - recommended cut-off point; SCID - structured clinical interview for DSM disorders; UAE - United Arab Emirates.

- unusually high estimates.

* Colour codes for adopted categories of changes in anxiety disorders likelihood: large - green, moderate - orange, small - red, negligible - white.

The most commonly used criterion was any anxiety disorder (7 out of 11 studies), followed by generalised anxiety disorder (4), with other types of anxiety disorders and sub-threshold symptomology, rarely being investigated. The majority of reference standards were DSM-IV criteria based, with an addition of one study that used the ICD-10 criteria, which are embraced in UK primary care practice.

Positive likelihood ratios were typically small and negative likelihood ratios were typically small to moderate. Large LRs- were more common than large LRs+, for both the HADS and GAD, i.e. the two tools were more accurate in ruling-out than ruling-in. Three studies had relatively small sample size (from 37 to 100). A validation study of Spanish version of PHQ-9 estimates resulted in an unusually high LR+ (Garcia-Campayo et al., 2010, LR+ 13 (95% CI 6.40, 27.0). Given that little information was provided about the execution of the reference test, a plausible reason for this discrepancy is unclear.

There were no directly comparable estimates of diagnostic accuracy across different criteria (i.e. based on the same instrument, cut-off, and reference standard). There were apparent differences in performance across specific anxiety disorders, with direct head-to-head comparisons available from three studies in adult primary care attendees (Kroenke et al., 2007, Bunevicius et al., 2007, Terluin et al., 2009). In one of these studies, the GAD-7 and -2 were compared across GAD, any anxiety disorder, panic disorder, PTSD and social phobia, where any anxiety disorder had the highest LRs+ (i.e. ruling-in accuracy) and GAD the highest LRs- (i.e. ruling-out accuracy) (Kroenke et al., 2007). In contrast, the HADS-A had the highest LR+ and LRs- for panic disorder as opposed to GAD, social phobia and any anxiety disorder (Bunevicius et al., 2007). Partially in line with this study, Terluin et al. (2009) have found for the HADS-A a higher LR- in patients with panic disorder/ agoraphobia/ social

phobia, than in patients with any anxiety disorder, but LRs+ were the same in both groups.

3.5.3.5 Responsiveness of GAD-7 and HADS-A

No eligible studies were found that presented original evidence on responsiveness for the GAD-7 in relevant populations. In one study with 273 patients with chronic pain recruited from a Swiss rehabilitation clinic, the HADS-A showed responsiveness (ES=0.35, SRM=0.45) to changes following four week in-patient interdisciplinary pain program (Angst et al., 2008).

3.5.4 Acceptability

The acceptability of selected depressive and anxiety measures in primary care was rarely investigated (see Table 3.14 overleaf for a summary of studies). Of these studies, one identified study considered psychiatric case identification questionnaires in UK general practice, and thus applies to anxiety measures (Wood et al., 2002). However, no data specific to anxiety measures could be identified. General acceptability of self-report measures and their underlying concepts and purposes are briefly summarised in the following sections - for both patients and health professionals.

Table 3.14 Summary of identified studies into acceptability of the concept of depression measures.

	Population /Country	N	Method	Condition of interest/ measure of interest	Study Aim
Coventry, 2011	Health professionals†, service users with diabetes and/or CHD, carers/UK	Stage I: 19 health professionals, 7 service users, 3 carers Stage II: 6 health professionals, 7 service users, 1 carer	Stage I: In-depth interview Stage II: Focus group	Depression/ any measures used in primary care	Barriers in management of depression in patients with long term conditions
Dowrick, 2009	GPs and patients/ UK	34 GPs 24 patients	Open ended, in-depth interviews	Depression/ assessment measures recommended by the QOF	Views on depression severity questionnaires incentivised in UK quality and outcomes framework
Rosemann, 2006	GPs, practice nurses and patients with OA/ Germany	20 GPs 20 practice nurses 20 patients with OA	Semi-structured interview with open-ended questions	Depression/ Depression questionnaires e.g. PHQ-9)	Problems and needs or improving primary care of osteoarthritis patients
Simpson, 2008	Primary care patients/UK	13	Semi-structured interviews	Depression/ PHQ-9	Patients' experiences of receiving collaborative care for depression
Wood, 2002	Primary care patients/ UK	127¥	20 focus groups	Psychiatric disorders/ an exemplar provided (GHQ-12)	Patients' attitudes to questionnaires for common mental health disorders

Note: GHQ - General Health questionnaire; GP - general practitioner; QOF - Quality and Outcomes Framework;
† - predominantly from primary care; ¥ - the number of groups which discussed each of the themes: tools validity (n=14), tools usefulness (n=20), the issue of confidentiality (n=20), perceived stigma (n=18).

3.5.4.1 Patients

Two qualitative studies detailed general attitudes of UK patients towards general self-report depression and self-report psychological measures [(Wood et al., 2002, Dowrick et al., 2009)]. Both studies concluded that the majority of patients were willing to complete questionnaires. The extent of equivocality differed across these two studies.

Wood et al. (2002) found that (particularly older) patients were likely to be ambivalent about psychiatric case identification questionnaires. Two main areas of concern were data confidentiality and perceived stigma (e.g. intrusive questions, fear of being judged). The majority of doubts were expressed in relationship to the tool validity (e.g. influence of temporary factors), but these concerns were likely to be stimulated by the provision of an exemplar questionnaire. Relative to other aspects of questionnaires, patients were more likely to raise advantages of their use (e.g. legitimising emotional problems).

In contrast, Dowrick et al. (2009) found that patients were in general less equivocal about self-report measures. They rarely raised the issue of their validity. Overall, they were likely to perceive questionnaires as more important than (or an important aide to) a clinician's judgment, although a holistic approach to interpretation was preferred. Patients were likely to perceive questionnaires as useful for management of symptoms (e.g. monitoring symptoms, improved self-awareness). This was confirmed by a small group of patients, receiving collaborative care for depression who found the PHQ-9 easy to understand and good at improving self-awareness (i.e. an acceptable way of monitoring progress of symptoms) [(Simpson et al., 2008)]. Overall, views expressed by patients in a small

number of studies suggest that questionnaires may be acceptable for primary care patients but this may differ according to their individual preferences.

3.5.4.2 Health care professionals

Following semi-structured interviews with 40 primary care health professionals, Rosemann et al. (2006) concluded that depression questionnaires (e.g. PHQ-9) are not used 'to reveal depression', but further details on their usefulness were not reported. Two small qualitative studies considered the views of healthcare professionals on self-report depression measures used in UK general practice [(Dowrick et al., 2009)] and in patients with long term conditions [(Coventry et al., 2011)]. In both studies health professionals expressed willingness to use questionnaires. They raised various advantages and disadvantages but overall they were likely to be equivocal.

GPs typically reported that self-report measures serve as a confirmation of their clinical judgment and are thus more useful for GPs with limited clinical experience. They suggested that in a general sense the scores needed to be interpreted in the context of patient information [(Dowrick et al., 2009)]. Likewise, a common view was that self-report measures may be inappropriate (i.e. reductionist or unsubtle) for patients with long term conditions, adding little to what can be offered by clinical judgment and continuity of care [(Coventry et al., 2011)].

Aside from GPs' valuing the questionnaires for their clinical judgement, a critical issue was surrounding the use of questionnaires in practice, i.e. acceptability. The difficulty was negotiating between a formal assessment according to the QOF and limited consultation time [(Dowrick et al., 2009)]. Similarly, a further important issue identified for primary care patients with long term conditions

was that higher work load was associated with using instruments, and using the recommended questionnaires in patients with poor English (Coventry et al., 2011).

Some GPs were positive about possible advantages of using depression questionnaires in practice (e.g. standardisation of care, matching intervention to severity) (Dowrick et al., 2009). Despite common resistance, formal assessment under the QOF recommendations was perceived by some health professionals as an improvement of management of depression in patients with long term conditions and a good experience for patients (Coventry et al., 2011). Overall, as health professionals were less likely to be positive about the purpose and usefulness of self-report measures of depression in UK primary care it seems that their acceptability may be limited.

3.6 DISCUSSION

3.6.1 Summary of key findings

For ease of comparison across properties and measures, a summary of findings is presented in Table 3.15.

Table 3.15 Summary of availability of evidence for the adapted evaluation framework.

	Depression symptom measures			Anxiety symptom measures	
Measurement property	**BDI**	**HADS-D**	**PHQ**	**GAD**	**HADS-A**
Feasibility					
- Administration burden	+	+	+	+	+
- Interpretability	?	?	?	?	?
- Accessibility	?	?	+	+	?
- Readability and comprehension	+	?	+	+	?
- Time to administer	+	+	+	+	+
Reliability					
- Cronbach's Alpha	+ (a)	+ (a, m)	+ (a, o)	+ (a)	+ (a, m)
- Item-total correlation	? (a)	? (a)	? (a, o)	? (a)	-
- Factor structure	? (a, m)	+ (a, m)	+ (a)	+ (a)	? (a)
- Intra-class correlation	-	+ (m)	+ (o)	+ (a)	-
- Test-retest correlation coefficient	-	-	+ (o, a)	+ (a)	-
Construct validity					
- Convergent validity	+ (o, m)	+ (a)	+ (a, o)	+ (a)	+ (a)
- Correlation with functional status	+ (a)	+ (m)	+ (m)	+ (a)	-
- 'Somatic bias'	+ (m)	+ (m)	-	-	+ (m)
Criterion validity					
- Ruling-out accuracy	? (m)	? (a, o, m)	+ (a, o)	+ (a)	+ (a); ?(o)
- Ruling-in accuracy	? (m)	? (a, o, m)	+ (a);?(o)	+ (a)	? (a, o, m)
Responsiveness	+ (m)	+ (a, m)	+ (a, o)	-	+ (m)
Acceptability					
- Patients			? (p)		
- Health professionals			? (p)		

Note: Identified evidence was classified as: + - supporting evidence; - - no data available; ? - doubtful (i.e. inadequate performance or weak evidence);
Kinds of study populations used: (a)- adult primary care adults, (o) primary care/community base adults, (m) - people with musculoskeletal complaints/ chronic pain, (p) - general primary care;
Marked in red - no UK data available.

There is some supporting evidence for most measurement properties for each of the measures. While feasibility and reliability appear well-supported, there

was less evidence for acceptability and criterion validity.

On the basis of this review it would be difficult to advocate the use of one particular measure over another. The key reasons for this are the lack of direct head-to-head comparisons within the same population, and the lack of large studies in the elderly and people with musculoskeletal complaints (or at least stratification within general adult primary care studies by age or presence of musculoskeletal conditions). A consequence of this paucity of evidence is the limited extent to which one can critically compare properties of each measure and concomitantly judge whether it is reasonable to assume that adequate performance in adult primary care will necessarily translate to older patients with OA. Assuming that data for primary care/community older adults is likely to reflect performance in older primary care attendees with OA, the PHQ-9 appears to perform relatively better than other measures. This echoes Nease and Malouin's (2003) conclusions. Of the anxiety measures compared in this review, the GAD seems to be promising, but currently lacks evidence in older and musculoskeletal complaint populations.

With specific reference to the HADS, this review suggests its adequacy as a research tool. It is short, easy to administer and has evidence for adequate internal-consistency and test-retest reliability. The HADS has adequate responsiveness to changes over time and thus is suitable for longitudinal data analyses. Being relatively free of 'somatic bias', it is unlikely to lead to score inflation or exaggeration of prevalence rates (therefore prevalence estimates based on HADS reported in chapter three, are least likely to be affected by OA or joint pain status).

A critical issue to consider for studies that aim for clinical applicability is how

to use the HADS to inform clinical practice. In the process of reaching a formal diagnosis, self-report severity measures act as an adjunct to clinical judgement based on symptoms count and functional impairment (NICE, 2009b), and indeed, GPs seem to share this view (Dowrick et al., 2009). Nevertheless, the choice of intervention is closely related to the classification of symptoms severity (NICE, 2009b). In the previous chapter, an approach based on classification of symptoms severity formed the basis for estimation of prevalence rates. Consequently, in the context of the current thesis, further analyses will continue to be focused on HADS defined classification of severity of symptoms.

Sources of bias in evidence

Aspects of psychometric and clinimetric properties and specific populations have been selectively investigated in the identified papers. This often appeared to be related to practical issues such as the original purpose of a specific tool and the year of development. Studies concerning the PHQ-9 and GAD-7 appeared more focused on feasibility, associations with functional outcomes and diagnostic accuracy as these tools are intended for diagnosing and monitoring primary care patients (Kroenke et al., 2001, Spitzer et al., 2006).

Whilst no systematic method of quality assessment was employed in this review, it has been observed that investigating criterion validity is particularly challenging. This could be related not only to limited access to relevant patients, but also the time-consuming nature of implementing 'gold' reference standards (i.e. clinical interview schedules). It is clear that selection and observer bias are difficult to avoid entirely, as through analyses of unusually low or high LRs several types of bias became apparent. For example, partial verification bias (e.g. Yeung et al.,

2008), disease progression bias (Axford et al., 2010), incorporation bias (e.g. Corapcioglu et al., 2004) and the execution of reference test was not described (Beck et al., 1997).

3.6.2 Strengths and limitations

Three key methodological decisions warrant critical considerations. These include a targeted approach to the included measures, taking a non-systematic approach to this review and the ability to fully specify the basis for judging 'adequate' performance.

It could be argued that the most commonly used measures (i.e. those listed in this review as used in OA or joint pain observational studies in general population/primary care adults) should be comprehensively reviewed. However, this would result in omitting the BDI-II and -PC and GAD-7 and -2, yet the BDI-II is advised for use in both UK primary care and for patients with rheumatologic problems. The GAD, in contrast, appears to be increasingly popular in UK primary care practice, but to date has not been compared against other anxiety specific measures. Consequently, the included questionnaires which were approved for primary care use, are of most relevance to primary care patients with OA.

The review was non-systematic, with one reviewer developing the search strategy, extracting data and judging the adequacy of the instruments performance. Furthermore there was no formal critical evaluation of the quality of the study designs. As discussed in the previous chapter, this suggests that this review is inevitably less robust and objective than a systematic review would be (i.e. it is potentially more predisposed to selection biases and errors). Nonetheless, the conducted search involved several approaches designed to identify all relevant studies, including searching systematically and including other data reviews and

their reference lists, followed by a broad search of electronic databases. By taking a non-systematic approach, it was possible to review a wider range of potentially relevant studies, and thus including evidence from primary care adults and older adults specifically, before seeking to critically compare this to patients with musculoskeletal disease.

This review sought to cover only selected measurement properties, omitting two clinimetric properties (content validity and cross-cultural validity). No comprehensive and pre-defined checklist for the quality assessment, was followed, such as the one developed by Bot et al. (2004). However, several psychometric properties were addressed and the review moved beyond this with considerations of clinical utility. This approach can be found in previous reviews of assessment measures (e.g. Meades & Ayers, 2010). In addition, to evaluate selected properties, specific indicative evidence has been-predefined and reported, and an effort was made to provide clear interpretation of criteria for judging 'adequate' performance.

3.6.3 Implications

Implications for clinical practice

Evidence of adequate measurement properties appears to be an expectation of any instrument to be used in clinical practice. The reviewed articles draw attention to the issues of reliability and validity of assessment measures being only part of the wider concern over the utility of self-report measures. This review touched on some important issues connected to utility such as acceptability, responsiveness, feasibility, and their usefulness of their application in practice. These require separate consideration in the context of improved

decision-making, patient outcome and cost-effectiveness, which were beyond the scope of this review. Overall, some gaps in performance of anxiety and depression questionnaires used in practice for older primary care attendees with OA are evident.

None of anxiety and depression measures reviewed are without disadvantages. The next question is: Can inadequate performance in one or more measurement properties affect their applicability and interpretation in practice for older adults with OA? In practice, the extent to which inadequacy on any of measurement properties is likely to affect applicability and interpretation of the results seems related to the degree to which the health professional rely on the measure with his interpretation and have 'adequate' skills to be independent of it. In these terms, inadequacy of one or more measurement properties is more likely to lead to misinterpretation in hands of health professionals, with less accurate clinical judgment, and subsequently, to be less applicable to older people with OA. Given that some GPs perceive questionnaire as 'more useful' (Dowrick et al., 2009), for less experienced healthcare professionals, the risk of misinterpretation is possible.

Research implications

Since specific anxiety disorders are uncommon in people with OA, future research may contribute to the current classificatory system of anxiety disorders, by validating a new category - *sub-threshold persistent anxiety disorder*. This can be inspired by criteria used for dysthymia (persistent sub-threshold depression symptoms present for at least 2 years) (APA, 2000, WHO, 1992).

The understanding of depression and anxiety symptoms in primary care populations with OA or joint pain is heavily reliant on the HADS data.

Consequently, it would be useful to investigate psychometric and clinimetric properties of the HADS and compare them with the BDI-II and -2, PHQ-9 and -2 and GAD-7 and -2 in a large sample of primary care patients with musculoskeletal pain, including sub-threshold forms of depressive and anxiety disorders.

3.7 CONCLUSIONS

The measurement properties of recommended depression and anxiety symptom measures have been reviewed. Although there is evidence to support some properties, a number of gaps in the evidence were identified. An understanding of how robust these measures are within the subpopulation of older primary care patients with OA is currently lacking, as are sufficient numbers of direct, head-to-head comparisons of measures within the same population. There are insufficient direct comparisons and studies within the subpopulations of interest to come to a confident recommendation for the most useful tool. However, the PHQ-9 – previously recommended by Nease and Malouin (2003) – has the potential to be a useful tool in older primary care attendees with OA.

Adequacy of the HADS was considered for estimating the longitudinal course of depressive and anxiety symptoms and detection of these problems in general practice. In the context of the former analyses, there is evidence to support HADS' stability whilst no changes in depressive and anxiety symptoms occur (as indicated by the adequate test-retest reliability) and responsiveness to occurring changes. Since the HADS does not include questions on aetiology of somatic symptoms of depression and anxiety, evidence to support this tool being relatively free from 'somatic bias' reduces a possibility of inflation of scores. The anxiety subscale seems to be more prone to this bias, than the depression

subscale. The critical issue for consideration in detection analyses is that the HADS is not a diagnostic tool and (as other reviewed measures) serves a relatively poor proxy for depressive and anxiety disorders. However, assuming that the HADS is suitable to assess the trajectories of symptoms and persistent symptoms are more easily identifiable over time than transient forms, valid conclusions can be drawn about detection of possible/definite anxiety and depression.

Chapter four: The course of anxiety and depression symptoms in older patients presenting to general practice with musculoskeletal pain

Part 1: Source and suitability of the PROG-RES dataset

4.1 INTRODUCTION

The primary objective of this chapter is to describe the data source used in subsequent chapters. Following a brief overview of secondary data analyses, this chapter describes the study design and sampling procedures used in the PROGnostic RESearch (PROG-RES) study. A brief introduction of the data collected and the characteristics of the study participants are presented. Suitability of the dataset, for the purposes of this thesis, is explored based on recommendations for secondary data analyses and with particular reference to the potential impact on analyses reported in chapters five, six and seven.

The purpose of chapters five, six and seven are to describe the course of anxiety and depression symptoms in older patients with osteoarthritis (associated person-related characteristics) and to assess detection of these coexisting psychological complaints in general practice. To achieve these goals, data from the PROG-RES study will be analysed.

4.2 BRIEF OVERVIEW OF SECONDARY DATA ANALYSIS

In the broadest sense, secondary data analyses can be defined as analyses conducted by a person not involved in either the study design or data collection (Boslaugh, 2007). According to Vartanian (p. 3, 2010) "[secondary data

analyses] *include any data that are examined to answer a research question other than the question(s) for which the data were initially collected*".

There is no single list of issues to consider whilst evaluating appropriateness of secondary datasets. Exploration of the source of data has been broadly regarded a major aspect of secondary data analyses (Boslaugh, 2007, Vartanian, 2010). According to Boslaugh (2007) for this purpose the following issues can be considered:

- *The original purpose of the study*: For instance, features of the targeted population or specific wording used might influence characteristics of data
- *Study design and sampling procedures:* The impacts of issues relevant to primary data such as time of data collection, methods of ascertainment, sampling procedures, response rate, or characteristics of the sample, are of concern to secondary data analysis
- *Data handling procedures:* This issue considers the practicality of data, including coding of missing data or specific variables

Vartanian (p.18 - 22, 2010) suggests that suitability of data can be evaluated by considering such issues as:

- *Feasibility,* including accessibility of data, required authorisations and knowledge needed for the statistical software to be used
- *Adequacy of the population* from which data was drawn
- *Generalisability* of sampling frame
- *Availability* and characteristic of dependent and independent variables
- *Adequacy of identifiers* for the target group and sample size of sub-groups

4.3 THE PROG-RESS COHORT: DESCRIPTIVE FEATURES

4.3.1 Design and setting

The PROG-RES study comprises a prospective cohort of consecutive, older people consulting their general practitioners about their musculoskeletal pain [(Mallen et al., 2006b)]. Adults aged 50 years and over consulting their general practitioner with a new or on-going episode of musculoskeletal pain were eligible for inclusion. The primary aim of the study was to evaluate the prognostic value of a brief assessment tool [(Mallen et al., 2006b)]. Excluded were patients with evidence of: traumatic injury; an acute swollen, red or hot joint; inflammatory arthropathy; or patients considered by GP to be vulnerable due to cognitive impairments or severe physical health problems. Patients were recruited from five Central Cheshire General Practices between September 2006 and March 2007 [(Mallen & Peat, 2008)]. Ethical approval was received for this study.

I had no role in the design or conduct of the PROG-RES study. The dataset is held by the Arthritis Research UK Primary Care Centre and I requested data to answer specific questions relevant to this thesis. The following sections will provide a brief overview of the design and conduct of the study, with a summary of key demographic findings provided to allow an understanding of the dataset used.

4.3.2 Data collection

A pilot study was completed between May and July 2006 at Kingsbridge Medical practice in Newcastle-under-Lyme. Following the implementation of minor amendments the study was conducted. Eligible participants triggered a specially designed electronic pop-up template by entering a predefined Read code [(Mallen et al., 2006b)]. If general practitioners decided that exclusion criteria did not apply, they

were requested to perform a brief assessment including six domains ascertained with seven questions and the consultation record was electronically 'stamped' (Mallen, 2009).

Following weekly electronic searches, stamped electronic records were identified and names and addresses of eligible patients were downloaded. Within one week of an electronically 'stamped' consultation (index), all eligible participants were sent a study pack including a letter from the general practice, an information hand-out, name and contact details of the principal investigator, a postal questionnaire and written consent for further contact and medical record examination. Non-respondents were sent a reminder postcard two weeks after the study pack.

Follow-up questionnaires were sent to all baseline respondents who consented for follow-up. Questionnaires were sent at 3, 6, 12 months, 2 and 3 years from the initial consultation date. Throughout the mailing process, weekly checks for patient deaths and departures from the general practices were conducted by the Research Network team. Details extracted from medical records included: date of index consultation, code assigned with a general practice and Read codes used in index consultation. Information was also extracted on prescriptions, comorbidities, referrals and other primary care use. Data entry, coding, cleaning and storage were detailed in the author thesis (Mallen, 2009) and the study protocol (Mallen et al., 2006b).

4.3.3 Response at baseline and follow-up

Information about participation rates from baseline to 6 months were extracted from the principal investigator's thesis. Details on 12 months, 2 and 3

years have been obtained from the custodian of the PROG-RES data. Sections below focus on time-points used in this thesis: baseline, 3, 6, and 12 months[1].

Baseline

Five practices participated in the study. Forty four GPs (Mallen, 2009) completed templates on 650 potential participants. Of the 646 eligible participants 502 responded to a baseline questionnaire (adjusted[2] response 77.7%) (Figure 4.1 overleaf).

Follow-up

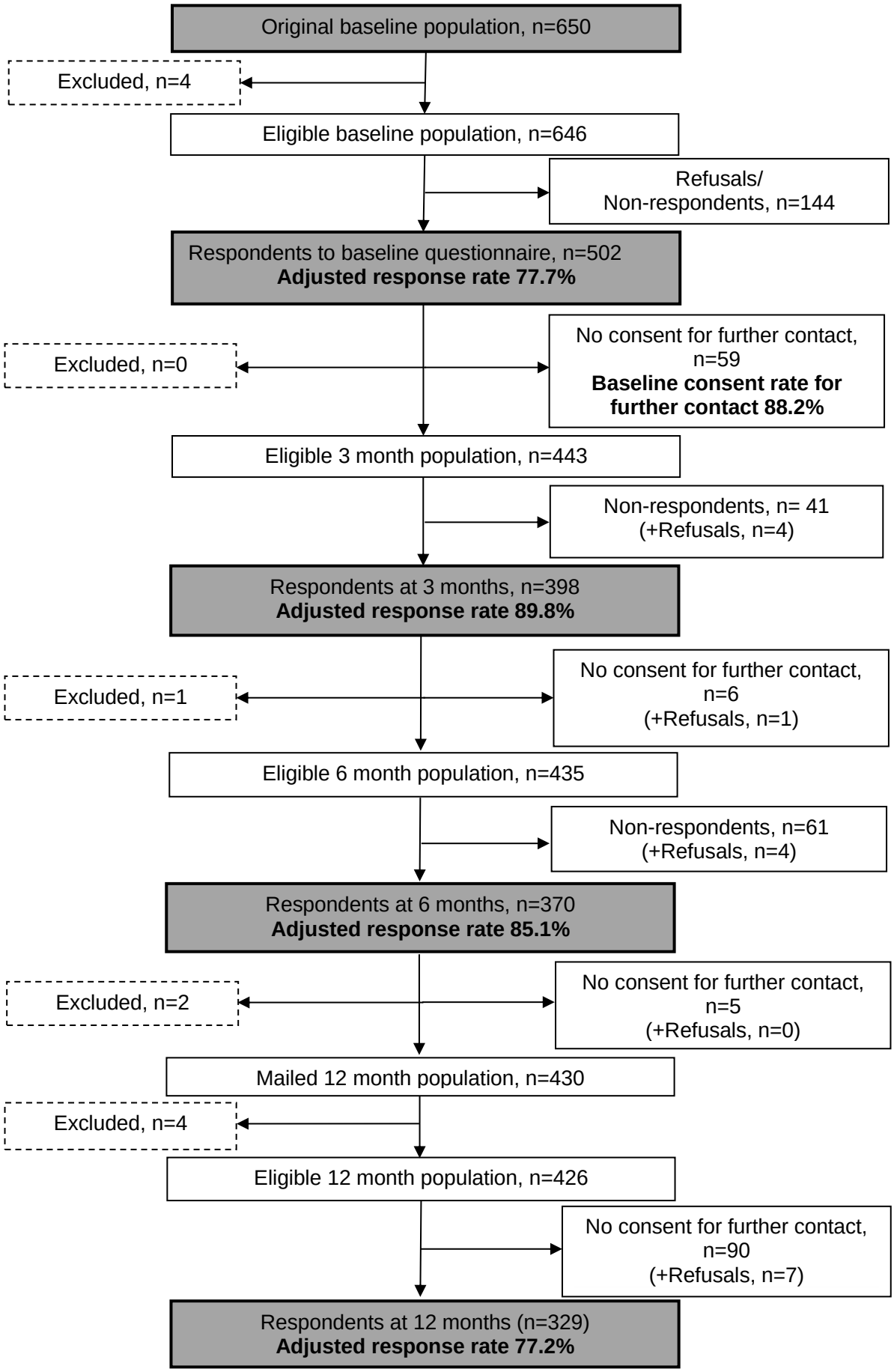
At the 3-month follow-up 443 participants were mailed a questionnaire, of which 389 returned completed questionnaires (adjusted response 89.8%). At the 6-month follow-up a questionnaire was emailed to 446 participants. Of these 370 completed their questionnaires (adjusted response 85.1%). The 12-month follow-up questionnaire was posted to 430 participants, of which 329 completed their questionnaires (adjusted response 77.2%) (Figure 4.1).

Consenters

Of the 502 respondents at baseline, 443 (88.2%) consented for follow-up. Permissions for medical record review were provided by 428 participants (a consent rate of 85.3%). In total, 403 participants consented at baseline for both follow-up and for medical record review (a consent rate of 80.3%).

[1] At 2 years n=250 mailed (214 respondents) and at 3 years n=244 mailed (202 respondents).
[2] An adjusted response rate refers to the response rate calculated with those who were excluded from the mailing process, due to death or died medical reasons or not living at the address to which the questionnaire was mailed.

Figure 4.1 Flowchart of 1-year follow-up responses to PROG-RES study.



4.3.4 Study participants

Sixty per cent (60%) of the study population were female. The mean age of participants was 63.1 years (SD 10.6) [(Mallen & Peat, 2008)].

4.4 THE PROG-RES COHORT: EVALUATION AS A SECONDARY DATA SOURCE FOR THIS THESIS

4.4.1 Generalisability of the sampling frame

Prevalence rates of 'mild or worse' anxiety symptoms (43.6%) and 'moderate or worse' anxiety symptoms (20.5%) were determined using HADS for the 502 baseline respondents. These were consistent with the pooled prevalence rates reported in chapter two (45.4% and 20.8% respectively). Prevalence rates of 'mild or worse' depression symptoms (27.9% vs. pooled 23.8%) and 'moderate or worse' depression symptoms (11.0% vs. pooled 14.6%) in the PROG-RES baseline respondents, were broadly comparable to the pooled estimates (chapter two). These findings support the generalisability of the cross-sectional characteristics of the main study outcomes. However, analyses of trajectories and detection rates could be affected by generalisability of the sample frame.

Statistical comparisons between the local and national populations were not conducted. In general, populations from all practices were likely to be white and have a higher than the national average Quality and Outcomes Framework (QOF) score [(Mallen, 2009)]. Consequently, ethnic minorities and practices with poorer quality of standards might be under-represented in the study. How this may impact on generalisability of the analyses of the trajectory of anxiety and depression in people with OA or joint pain is unclear. Both race and socio-economic status has been shown to affect a GP's perception of patients, which can affect consultation

behaviour [van Ryn & Burke 2000]. Assuming that quality of doctors' care for patients is important to depression outcome [Croghan et al., 2006], detection rates of psychological problems can be expected to be positively associated with increasing QOF scores.

The participants suffered severely disabling pain, more severe than reported in previous observational research [Mallen, 2009]. This was unlikely to limit generalisability of the sample to UK primary care patients, as evidence suggests that UK primary care patients reporting joint pain are more likely to suffer severely disabling pain than mild or moderately disabling pain [Thorstensson et al., 2009]. However, an obvious consequence of this observation is that, it can be expected that at the time of the study PROG-RES participants were already in the process of adjusting to pain. Subsequently, it is possible that many of them could have experienced symptoms of depression and anxiety prior the study.

4.4.2 The risk of response bias and loss of precision due to attrition, non-consent to follow-up and medical record review

Response bias

Non-response bias is a systematic difference in characteristics of those who responded when compared with those who did not respond, which can reduce the representativeness of the contacted sample [Jooste et al., 1990]. High response rates are believed to reduce the chance of response bias [Singleton & Straits, 2005]. According to Bowling (2002) response rates of 75.0% and above could be considered good, and thus the response rate of over 77.7% appears to reduce the chance of response bias. However, in the PROG-RES women appeared to be more likely to respond than men, but statistical significance was not tested for [Mallen, 2009]. In accordance with previous research [Dunn et al., 2004], response was the highest in age

group 70-79 years and the lowest for younger men (50-59 years) and older women (80 years or older) [Mallen, 2009]. The precise impact this could have on the trajectories and detection rates presented in subsequent chapters is unclear.

Consent bias

Participants in the oldest age group (80 years or above) were less likely to consent for follow-up (e.g. 80 years or above: 74.0% vs. 50-59 years: 91.5%). Women in the oldest group were less likely to be represented in medical records (e.g. females aged 80 years or above: 79.4% vs. males aged 70-79 years: 94.1%). As indicated by previous research, the decreased likelihood to consent for follow-up in older age is a common problem [Dunn et al., 2004]. This is most likely to be due to unwillingness to participate- a common reason for lack of consent in older medical patients [Bakke et al., 1990]. Given that older people are less likely to consult their GPs for depressive and anxiety symptoms [RCGP, 2006], the results of medical record review analyses could be influenced by under-representation of the oldest adults. Comparisons of non-consenters and consenters for medical record review and further contact suggests that those consenting to the latter were more likely to be older, widowed, have poorer physical health and more depression symptoms. Statistical significance was not tested. Overall, there is some limited evidence for consent bias by age in the PROG-RES study [Mallen, 2009]. Nevertheless, comparisons of baseline respondents and consenters for further contact and medical record review (conducted by the principal investigator of the PROG-RES study [Mallen, 2009]) demonstrated that consent for further contact at baseline was similar across age, gender, general health, psychological and pain characteristic. These suggest that any consent bias was unlikely to be substantial.

Attrition bias

Attrition affects longitudinal investigations and occurs when participants are lost to follow-up (Porta, 2008). The effect of drop-out was further ascertained by the PROG-RES principal investigator using variables explored in consent analyses. As the current study looks at one year trajectories only, this time frame sufficed. Comparisons were made between baseline consenters for medical record review and further contact, respondents to the 6-month follow-up and non-respondents at six months. The observed pattern appears to be broadly comparable to that found for consenting, except for age. Although typically older adults are less likely to successfully be followed-up (Chatfield et al., 2005), in the PROG-RES study, participants aged 70 years or older were more likely to respond at follow-up. According to Mallen (2009) this could be due to interest in the study topic and short intervals between follow-ups. The exact impact of this possible attrition bias by age on the current study is unclear.

4.4.3 Case definition for musculoskeletal pain

The most common reasons for consultations reported on the baseline questionnaire included low back pain, followed by shoulder and knee pain (Mallen, 2009). On the baseline manikin, the knee was the commonest pain location, followed by low back and shoulder pain. In total, of the 502 baseline participants 454 (90.4%) had pain at more than one location (Mallen, 2009) (Table 4.1 overleaf). Although all participants were older adults and joint pain was frequently reported, clearly musculoskeletal pain cannot be assumed to be all related to OA.

Table 4.1 Pain location, by index consultation and manikin data.

Anatomical site	Pain location	
	Consulting n (%)	Manikin n (%)
Ankle/foot	99 (19.7)	173 (34.5)
Low back	148 (29.5)	232 (46.2)
Elbow	31 (6.2)	102 (20.3)
Hip	90 (17.9)	209 (41.6)
Knee	127 (25.3)	250 (49.8)
Neck	75 (14.9)	148 (29.4)
Shoulder	133 (26.5)	211 (42.0)
Wrist/hand	68 (13.5)	127 (25.3)
Other	82 (16.3)	-

Note: Numbers not equal 502, because some patients consulted with and shaded pain at more than one anatomical site.

A selective sample of patients with an OA diagnosis could be identified through reviewing medical records. However, this would lead to the loss of power of analyses due to a reduced sample size. Furthermore, the use all forms of musculoskeletal pain is supported with ample evidence suggesting that the specific diagnosis of OA can pose difficulties for GPs (Bedson et al., 2005, Bierma-Zeinstra et al., 2000). Diagnostic labels for knee OA seem to be used with caution by primary care practitioners in the UK, as diagnosis of knee OA was found not to be an initial response to the presentation of knee symptoms (on average taking 10 years to be allocated a formal knee OA diagnosis) (Bedson et al., 2005). Likewise, a specific diagnosis was registered for only 32% of the 400 older general practice patients (20 per individual GP) with new hip problems in the Netherlands (Bierma-Zeinstra et al., 2000). The number of patients who received a hip OA diagnosis varied from 5% to 50% (Bierma-Zeinstra et al., 2000).

Diagnostic labels of OA may also be used inaccurately by GPs. For example, the diagnostic concordance between GP diagnosis of knee OA and ACR clinical classification was found to be fair (Kappa (k) =0.28, 95% CI -0.01, 0.56) in general practice attendees in the UK (Peat et al., 2005). Likewise, levels of agreement between rheumatologists and referring primary care practitioners in patients with musculoskeletal disease in Spain, were found to be moderate for peripheral OA (k=0.48, 95% CI 0.38, 0.57), knee pain (k=0.40, 95% CI 0.29, 0.59) and muscular pain diagnosis (k=0.15, 95% CI 0.10, 0.20) (Candelas et al., 2010). Difficulties with diagnosis can be partially related to the lack of valid standardisation of diagnosis (Bierma-Zeinstra et al., 2000). Since radiographs are infrequently used in initial diagnosis of a patient with probable OA (as recommended by NICE (2008)) and nearly half of patients with radiographic evidence of OA are asymptomatic, the diagnosis of OA is most often reliant on the GP's clinical skills (Altman et al., 2009).

Inclusion of patients with spinal problems seems justified on the basis of common coexistence of spinal pain and peripheral joint. For example, Suri et al. (2010) reported that over half of all patients with symptomatic tibiofemoral knee OA had pain in low back. Low back pain was found correlated with higher scores on the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) (Wolfe, 1999). It can be argued that this correlation may simply be a marker for individuals with a greater propensity to pain states (Natvig et al., 2001). Some researchers believe that low back pain is biomechanically linked to knee pain via the so-called 'knee-spine syndrome' (Tsuji et al., 2002). Indeed, Suri (2010) found that pain in low back and ipsilateral foot pain were significantly and independently associated with higher WOMAC knee pain scores, and concluded that join affects nearby joints both above and below in the kinetic chain.

Lumbar spinal facet joints are suggested to be a source of low back and lower extremity pain, with an on-going debate over this idea in the medical literature [Borenstein, 2004, Kalichman et al., 2008]. No significant relationships were found between degenerative changes in the lumbar spine facet joints and symptomatic low back pain [Schwarzer et al., 1994, Kalichman et al., 2008]. However, supporting evidence (for lumbar spinal facet joints being a source of low back pain) come predominantly from decreased back pain following intra-articular or peri-articular joint injections [Lewinnek & Warfield, 1986, Schwarzer et al., 1994]. Furthermore, many of the therapies for osteoarthritis have generalised effects throughout the body, i.e. therapies found to be effective for the appendicular skeleton are also effective for axial skeletal disease [Dieppe & Brandt, 2003].

4.4.4 Ascertainment of anxiety and depression symptoms

Suitability of HADS as an instrument

In the PROG-RES study, depressive and anxiety symptoms were ascertained with the HADS. The suitability of this questionnaire, in the context of this thesis, was considered in chapter three, where the absolute and relative performance of the HADS was evaluated. HADS is a standardised and feasible longitudinal research tool and widely used to assess severity of depressive and anxiety symptoms, but serves a poor proxy for depressive/anxiety disorders (see a summary on pages 125-126 and conclusions on page 130). A cut-off score ≥8, inclusive of mild to severe symptoms is of clinical relevance to primary care [NICE, 2009b]. Three recommended score range [Zigmond & Snaith, 1983] and their severity classification [Kendrick et al., 2009] include; 0-7; none, 8-10; mild symptoms, ≥11; moderate to severe. Previous analyses conducted for PROG-RES data, however,

also included cut-off score ≥15 being indicative of severe symptoms (Mallen & Peat, 2008).

Suitability of the time frame

The PROG-RES study provides six repeated HADS measures. HADS scores collected at baseline, 3, 6 and 12 months were used in identification of the trajectories of depressive and anxiety symptoms in primary care attendees consulting their GPs about musculoskeletal pain. Two and three year follow-up HADS data was excluded from the analyses presented in this thesis due to the reduced numbers of respondents. The choice of the time frame and frequency of repeated measures still provided multiple repeated measures which according to Harris et al. (2006) can be useful to depict relapsing of symptoms.

Coexistence of depressive and anxiety symptoms

Table 4.2 overleaf shows that 47.0% PROG-RES respondents had no anxiety or depression symptoms. 'Mild or worse' depression symptoms without anxiety symptoms were present in 7.0% of patients. This pattern was broadly comparable across four time points. The statistical implication of this pattern for this thesis is a need to increase a sample size by independently focusing on anxiety and depression symptoms, regardless of coexisting depressive/anxiety symptoms. This decision is supported by clinical reasons as primary care patients frequently experience coexisting depressive and anxiety symptoms (Mitchell et al., 2011).

Table 4.2 Frequency of co-occurrence of mild or worse anxiety and/or depression symptoms (HADS score ≥8).

	Baseline N=502	Month 3 N=398	Month 6 N=370	Month 12 N=329
	n (%)	n (%)	n (%)	n (%)
Neither	236 (47.0)	201 (50.5)	191 (51.6)	168 (51.1)
Anxiety and depression	102 (20.3)	84 (21.1)	70 (18.9)	58 (17.6)
Only anxiety	117 (23.3)	80 (20.1)	76 (20.5)	78 (23.7)
Only depression	35 (7.0)	21 (5.3)	16 (4.3)	20 (6.1)

Note: Percentage is not equal 100% due to missing scores.

Distribution of HADS data

Means, standard deviations and the result of tests for asymmetry are presented in Table 4.3. Together with visual examination, Kolmogorov-Smirnov and Shapiro-Wilk tests of normality (that yielded significant results at levels $p<0.0001$ for both anxiety and depression symptoms at each time point) suggest an asymmetry from the normal distribution of HADS data. A consequence of this non-normality is that it limits the use of parametric tests.

Table 4.3 Asymmetry of HAD depression and anxiety scores at 4 time point.

Time point	n	Mean	Standard deviation	Skewness		Kurtosis	
				Statistic	SE	Statistic	SE
HADS depression score (0-21)							
Baseline	502	7.21	4.112	0.373	0.110	-0.178	0.220
3 m	398	7.08	4.094	0.364	0.124	-0.249	0.247
6m	370	6.82	4.295	0.338	0.129	-0.481	0.257
12m	329	6.64	4.183	0.350	0.135	-0.416	0.270
HADS anxiety score (0-21)							
Baseline	502	5.67	3.633	0.748	0.110	0.208	0.219
3 m	398	5.47	3.654	0.588	0.123	-0.119	0.246
6m	370	5.11	3.923	0.928	0.129	0.663	0.257
12m	329	5.05	3.753	0.600	0.135	-0.281	0.270

Note: HADS - Hospital Anxiety and Depression Scale; SE - standard error.

4.4.5 Availability of covariate data

The PROG-RES study offers access to a range of baseline data including general demographics, physical health status, pain characteristics, psychological and social factors. Two basic demographic characteristics were included, namely age and gender. Lee and Mercurio-Riley's (2009) conceptual framework of factors found to be associated with psychological adjustment to chronic pain (mainly originating in musculoskeletal pain) (see Figure 1.1 on page 15), was used to select covariate data. Selected factors were based on principles of: availability in the PROG-RES study; parsimony; and limited instability of a factor (to reduce a confounding impact of transient states). Data on *intrapersonal factors* (e.g. self-efficacy, personality, readiness) and *stress factors* (psychopathologies, movement avoidance, and pain-related stressor) was not available from the PROG-RES questionnaire data. As the selected frameworks are operational this is not detrimental to the result of analyses.

Demographic characteristics

Age ("date of birth") and gender ("male/female") were assessed using standard, single questions. The method of grouping age was categorised as per the primary analyses of this dataset (50-59, 60-69 and 70 years or above) (Mallen et al., 2007). This way consistency was maintained, without compromising the size of age groups.

Pain condition (Lee and Mercurio-Riley's, 2009)

The type of pain conditions included in the PROG-RES study were non-inflammatory musculoskeletal pain, including OA. No further distinction was made

by diagnosis, but information on the pain location and severity was available. Anatomical pain location was assessed with a manikin. This method involves using diagrams of front and back views of a body manikin on which participants are asked to shade the location of their pain with reference to pain experience in the past four weeks. There is evidence for the tool's inter-rater reliability (Lacey et al., 2005), but the validity warrants further investigation. Shaded manikins can also be used to assess the extent of pain (Lacey et al., 2005), which can be determined by using established criteria such as separating a manikin into 44 mutually exclusive anatomical areas and counting the number of pain sites (Lewis et al., 2002). This method has been used in this thesis.

Pain severity was assessed using the 7-item Chronic Pain Grade (CPG) scale. This measure is a composite of pain intensity, disability days and interference with activities, which using scoring rules grades pain severity into 4 hierarchical classes (Von Korff et al., 1992). The measure has good reliability and convergent and construct validity (Smith et al., 1997). Only selected CPG items were included in the analyses and full details are described in the next section.

Functional dependence (Lee and Mercurio-Riley's, 2009)

The PROG-RES study collected data on pain-related interference with social, work, and daily activities, assessed with three items on the Chronic Pain Grade (Von Korff et al., 1992). Each of these questions refers to the past three months and are graded 0 ("no interference") to 10 ("unable to carry on any activities"). The idea of using questions on pain interference with everyday life activities independently as a proxy for pain interference with activities is not novel in musculoskeletal pain research, with these items being previously extracted from

general measures such as the MOS SF-12 [Thomas et al., 2004]. Scores computed from these CPG pain-interference items have not yet been directly tested for validity and reliability in persons with OA. However, the CPG pain-interference score showed increasing correlation with mental health, general health, emotional problems, energy, physical functioning, physical problems, social functioning and pain subscales on SF-36, at baseline in UK general practice patients [Smith et al., 1997]. The CPG, which incorporates these items, has shown increasing associations with unemployment rate, functional limitation, depression, self-rated health, use of opiates, and pain-related doctor visits at a 1-year follow-up in primary care patients with chronic pain, including back pain [Von Korff et al., 1992].

Stress processing factors (Lee and Mercurio-Riley's, 2009)

Information on participant's use of selected pain coping strategies was available in the PROG-RES from an abbreviated version of the 2-item version of the Coping Strategies Questionnaire (CSQ) [Jensen et al., 2003]. The selected coping strategies were: catastrophising, increased behavioural activities; coping self-statements, and ignoring pain. The frequency with which participants report using each coping strategy is assessed by two items (8 items in total in the PROG-RES survey instrument) with response options indicated on a 7-point numerical rating scale ranging from 0 ("never do") through 3 ("sometimes do") to 6 ("always do that"), without a specified time frame. For the purpose of analyses conducted in this thesis the highest tertiles were calculated for each of the four pain coping strategies and used as cut-off points. This was based on the assumption that people are likely to try several coping strategies, with those strategies which are used often more likely to persist over time [Craighead & Nemeroff, 2004]. Highest tertile

points estimated for the PROG-RES dataset, ranged from 4 to 5 across individual coping strategies, and thus, overlapped with answer options – between "sometimes do that" and "always do that" – suggestive of strategies being used often. Internal reliability and validity have been confirmed for each coping strategy (Rosenstiel & Keefe, 1983, Jensen et al., 2003).

Socio-ecological factors (Lee and Mercurio-Riley's, 2009)

Six socio-ecological factors were assessed in the PROG-RES: living arrangements; marital status; availability of emotional and instrumental support; occupation and employment status. Living arrangements were assessed by asking participants whether they were living alone ("yes/no"), a question that has been widely used in research (e.g. Thomas et al., 2004, Iliffe et al. 2009). Marital status was assessed using a single question (Thomas et al., 2004), with six answer options available ("married/ single/ divorced/ widowed/ separated/ cohabiting"). As in previous research (Mallen et al., 2007), for the purpose of analyses described in subsequent chapters, marital status was dichotomised into married/cohabiting vs. not married/cohabiting. The availability of instrumental and emotional support were each assessed using one question asking participants if they had anyone providing emotional support and extra help ("no/yes/no need") (Krumholz et al., 1998). Responses were dichotomised into no vs. yes/no need.

Two free text questions on current or most recent job titles were utilised to establish socio-economic classes. As in previous research (Jordan et al., 2008) the National Statistics Socio-Economic Classification (NSS-EC) (Office for National Statistics, 2010) was used for this purpose, with socio-economic status divided into four classes (managerial or professional/ intermediate/ routine or manual/ other (e.g.

house wives and never employed)) (Office for National Statistics, 2010, Jordan et al., 2008). Responses were grouped into manual/routine vs. not manual/routine for the purpose on statistical analysis. Current employment status was assessed in the PROG-RES by asking participants to classify their employment status as one of the following: “Employed/ not working due to ill health/ retired/ housewife/ unemployed/other”. Employment status was excluded from analyses conducted in this thesis, due to a high proportion of participants being retired.

4.4.6 Completeness of the key data

The levels of missing data amongst scales and items used in this thesis can be found in Table 4.4 overleaf. Levels of missing main outcome data were low, with a small decrease at the 6-month follow up point. Low levels of missing data were found for all covariate data, except for coping strategies. High levels of missing CSQ data (17.9%) have been previously reported in a community-dwelling sample of older adults (Kovacs et al., 2008), which the authors attributed to using the length of the survey. In the PROG-RES study, the CSQ appeared after the HADS. This suggests that the content of the CSQ questions could be perceived by some participants as sensitive. This finding is interesting, as it might suggest that in general the investigated sample was receptive to reporting depressive and anxiety symptoms, but less acceptant of considering their role in management of pain.

Table 4.4 Levels of missing data amongst items and scales at baseline, 3-month, 6-month and 12-month follow-ups.

Measure: Variable	PROG-RES at baseline (n=502) %	PROGR-RES at 3- month follow-up (n=398) %	PROG-RES at 6-month follow-up (n=370) %	PROG-RES at 12-motnh follow-up (n=329) %
HADS-D: depression symptoms	1.2	1.5	3.5	1.2
HADS-A: anxiety symptoms	2.2	2.3	3.2	1.5
HADS*: depressive and anxiety symptoms	2.4	3.2	4.6	1.6
Pain manikin: number of pain sites	0.0	N/A	N/A	N/A
CPG: Interference with daily activities	4.0	N/A	N/A	N/A
CPG: Interference with work	3.8	N/A	N/A	N/A
CPG: Interference with social activities	3.8	N/A	N/A	N/A
CSQ-2: Catastrophising	9.0	N/A	N/A	N/A
CSQ-2: Coping self-statements	8.0	N/A	N/A	N/A
CSQ-2: Coping by ignoring pain	10.4	N/A	N/A	N/A
CSQ-2: Coping by increased behavioural activities	9.4	N/A	N/A	N/A
1 item: Living arrangement	0.4	N/A	N/A	N/A
1 item: Marital status	0.8	N/A	N/A	N/A
1 item: Instrumental support	0.6	N/A	N/A	N/A
1 item: Emotional support	1.0	N/A	N/A	N/A
Occupation: socio-economic class	2.0	N/A	N/A	N/A

Note: CPG - Chronic Pain Grade; CSQ - Coping Strategies Questionnaire; HADS - Hospital Anxiety and Depression Scale; N/A - data not used in this thesis;
* - missing either or both scores.

4.4.7 Size of cohort and statistical power for analyses

The 443 participant's consenting to follow-up formed the basis for the analyses described in chapters five, six and seven. There is no minimum sample size specified for analyses of trajectories or medical records data. Nandi et al. (2009) identified 29 studies that assessed heterogeneity in the trajectories of depression and anxiety, with sample sizes of the identified studies ranging from 157 to 11559. There are examples of comparable sample sizes used in identifying

trajectories of depression and anxiety in other medical populations, for instance in 475 patients after myocardial infarction (Kaptein et al., 2006) and 398 patients with breast cancer (Dunn et al., 2011). There is a study in which medical records of 2294 patients were used (Kendrick et al., 2009). However, accessing medical records data appears challenging, consequently it is not uncommon to use samples ranging between 150 and 200 (Dowrick & Buchan, 1995, Cully et al., 2009) or even less than 100 (Licht-Strunk et al., 2009a).

Subsequently, the sample size of the PROG-RES appears relatively small, but not exceptional, and thus adequate for the planned analyses. This will however, have implications for statistical analyses, including: the number of included covariates and categorisation of responses; the choice of the minimum required sample size for trajectories; as well as the choice of statistical tests for analyses of detection (where the sample size is expected to further decrease).

4.5 CONCLUSIONS

The PROG-RES data appears applicable to UK primary care patients with symptomatic OA. This has been achieved by recruiting older consecutive primary care consulters with non-inflammatory musculoskeletal pain including OA. As a longitudinal data source it enables the current research to meet objectives set out for the next chapters. The available time frame and multiple repeated measures allow investigation of the relapsing nature of psychological distress symptoms without compromising the power of analyses. The combination of self-report and linked medical record data is an advantage.

The dataset enables access to the outcome variable and relevant covariate data, including the HADS recommended for primary care. The methods of

ascertainment have been previously used in research and are mostly standardised. In general, response, consent and follow-up rates for the investigated time were good and the key data is well-completed, with unlikely substantial response, consent and attrition bias.

It is unclear to what extent the PROG-RES data is generalisable to people from ethnic minorities, patients registered with practice with below average QOF scores, and with mildly to moderately disabling pain. Participants with non-OA musculoskeletal pain were included in analyses conduced for the purposes of this thesis. These issues are critical considerations, rather than limitation *per se*. The relatively small sample size is a limitation of the PROG-RES dataset. This will be considered in further methodological decisions.

As the PROG-RES study was not designed to meet the purposes set out for this thesis, it is not free of issues inherent to secondary data analyses, such as the choice of measures and availability of covariates. As described in chapter three, the HADS diagnostic use is limited without support from clinical expertise. Investigating the support and relationships with significant others might also benefit from taking a qualitative rather than a quantitative approach. The dataset does not allow for investigating *intrapersonal* and *stress factors*. Clinical factors, such as the effect of depression or anxiety treatments, belief in powerful others (e.g. doctors and medical treatment), patient attitude to recognition of symptoms and diagnosis of depression and anxiety and perceived need for help cannot be explored.

Chapter five: The course of anxiety and depression symptoms in older patients presenting to general practice with musculoskeletal pain

Part 2: A latent class growth analysis

5.1 INTRODUCTION

Chapter two identified that anxiety and depression symptoms commonly coexist with OA or joint pain in older people. To date, much of this information comes from cross sectional general population surveys with less being known about the course of anxiety and depression symptoms over time. This chapter considers the course of depression and anxiety symptoms (determined using the HADS questionnaire) over a 12-month period.

5.2 METHODOLOGICAL BACKGROUND

5.2.1 Variable-centred vs. person-centred methods for identifying the course of symptoms

Muthén and Muthén (2000) highlight the differences between person-centred and variable-centred approaches. The latter approach uses methods such as prevalence rates, structural equation modelling, regression or factor analysis. These types of analyses aim to explore the relationships between variables by identifying predictors of outcome variables or relationships between dependent and independent variables.

In contrast, person-centred approaches such as cluster analysis, latent class analysis, and finite mixture modelling, explore the relationships among people by classifying them into distinct groups or categories based on individual

response patterns. The objective of this approach is to extract information about inter-individual differences in intra-individual changes over time (Nesselroade, 1988). The analysis presented in this chapter will focus on person-centred longitudinal approaches for two main reasons: a) It can provide information about individual patient outcomes over time and b) It can separate the effect of changes experienced over multiple time points by an individual patient from between-subject differences at baseline (cohort effect). Person-centred analyses will be complemented with cross-sectional analyses (i.e. changes in the rate of symptom reporting), to enhance comparability with previous research.

5.2.2 Choice of a person-centred method of identifying symptom trajectories

Three methods specifically tailored for modelling longitudinal data, which allow the use of a single outcome variable at multiple time points to define discrete subgroups were compared: latent class growth analyses (LCGA), growth mixture modelling (GMM) and longitudinal latent class analyses (LLCA). Critical considerations were made between these techniques, including the work of Croudace et al. (2003), Feldman et al. (2009) and Peng (2011). Table 5.1 overleaf shows a summary of key considerations made during comparisons.

Table 5.1 Key critical considerations made to choose the method of modelling longitudinal data.

Critical considerations	Longitudinal latent class analyses (LLCA)	Latent class growth analyses (LCGA)	Growth mixture modelling (GMM)
What is modelled?	Patterns of state across time	Patterns of change over time on a given variable	Patterns of change over time on a given variable
Does it capture changes across multiple time points?	Yes	Yes	Yes
Is time order accounted?	No	Yes	Yes
Is within-class variation not allowed?	Yes	Yes	No
Is user-friendly statistical software available?	Yes	Yes	No
Are standardised methods of assessing goodness of fit of the model available?	Yes	Yes	Yes
Other critical considerations?	Comparable trajectories to LCGA with covariates. An increased risk of violation of local independence. Data is required to be normally distributed.	Serves as a good starting point for conducting GMM. Data is required to be normally distributed.	Difficult to interpret. Data is required to be normally distributed. Exploration for within-class predictors is enabled.

Longitudinal class growth analysis (LCGA) was selected for the following key methodological reasons: a) it defines distinct groups of individuals who share patterns of progress b) it can be used for multiple re-measurement points and c) time order is allowed. LCGA also appeared to be feasible due to its relatively easy

interpretation, as variation within a cluster is not allowed and user-friendly statistical software is accessible (Latent GOLD).

Statistical justification for performing LCGA

The process of estimating symptoms over time, using group based modelling, has been described as "cumbersome" (Delucchi et al., 2004). In addition to clinical benefits, the need for using a person-centred method of analyses is considered justified, if progress of symptoms is expected to vary (Laursen & Hoff, 2006). In the investigated sample this could be judged by establishing a number of observed individual patterns of symptoms with a frequency of occurrence ≥1% (Peng , 2011).

Details of the LCGA model for binary data

LCGA, referred to as a group based modelling (Nagin, 2005), allows for grouping of individual differences in longitudinal trajectories into a finite number of latent clusters. It is a semi-parametric method (Nagin, 2005), with parameters describing individual level trajectories that are assumed to be distributed according to a multivariate normal distribution (Titterington et al., 1985). If this criterion is met, each cluster can be an estimate of true population parameters. If the data is not normally distributed, it should be accommodated by non-distribution based methods of model estimation, including categorical data (Bauer & Curran, 2003).

For binary data LCGA assumes that variance in a sequence of a repeated binary observed variable C (in this case depression or anxiety) measured at t time points for a person i, can be explained by a latent categorical variable X (latent clusters) with j clusters and j=1,2,...J. The probability of having depression/anxiety

($C_{it} = 1$) at each time point t=1,2,....T is estimated for each cluster. The estimation of having depression/anxiety ($C_{it} = 1$), given being in cluster j, is made using the binary logit distribution (Nagin, p. 35, 2005). The model implemented in this study is embodied in the following equation:

$$P(C_{it} = 1|X = \mathrm{j}) = \frac{e^{\alpha^j + \beta_1^j \eta_t + \beta_2^j \eta_t^2}}{1 + e^{\alpha^j + \beta_1^j \eta_t + \beta_2^j \eta_t^2}}$$

Where $\mathrm{C_{it}}$ represents depression/anxiety for a patient i at t time points (t =1,2,3,4), X is a latent binary variable with J categories (X=j, j=1,2,...J). For each cluster j, LCGA estimate the mean growth curve which is a function of time and contains growth factors such as intercept (α^{j}) and linear slope (β_1^{j}) and quadratic slope (β_2^{j}) parameters; and η_{t} is the factor loading of time with specified time order of observed variables. In addition to the time-order effect, LCGA assumes that the variance and covariance estimates for the growth factors with each cluster are fixed and equal zero (Nagin, 2005). In other words growth factors of each individual i are equal to the mean growth factor for all individuals in cluster j (Peng, 2011).

Classification of cluster membership

A commonly used method for assessing cluster membership is the posterior membership probability derived from Baye's theorem (see Peng, p. 41, 2011). Allocation of anxiety/depressive outcome to cluster j is based on maximum posterior probability with the sum of all posterior probabilities for each individual equal 1. In other words, following this assignment depressive/anxiety outcome is allocated to j-cluster where posterior probability is the highest.

Choosing the optimal model: judging goodness of fit

There is not one single criterion for judging model fit. Instead, several indices of goodness of fit are available, the ones used in this study included:

- **Log-likelihood (LL)** test was conducted with Latent GOLD to assess the differences in the log-likelihood between neighbouring LCGA models (Peng, 2011). The LL difference is denoted as log-likelihood ratio (LR) and is defined as $-2(\mathrm{LL_j} - \mathrm{LL_{(j-1)}})$, where $\mathrm{LL_j}$ refers to the log value of the maximised likelihood of the j-cluster model and $\mathrm{LL_{(j-1)}}$ is the log value for the j-1 cluster's maximised likelihood (Nylund et al., 2007). The change in the log-likelihood statistics from (j-1) to j-cluster model can be explored using a graphic representation, with a flattened pattern being indicative of little improvement in model fit by adding j+1 cluster.
- **Bayesian Information Criterion (BIC)** was also used (Schwarz, 1978, Kass & Raftery, 1995). The BIC for the log-likelihood for the model with j-cluster ($\mathrm{LL_j}$), and is denoted as $-2\mathrm{LL_j} + \mathrm{Npar}\,(\log \mathrm{n})$, where Npar is the number of free parameters in the j-cluster model and n is the sample size. The number of clusters is increased until BIC minimum is found. With the lowest value of BIC indicating the optimal number of clusters (Nylund et al., 2007).
- To obtain an additional indicator of improvement of fit by adding a class (j+1), a bootstrap estimate of the p-value related to the difference in log-likelihood value (-2LL-difference) between two models with j and j-1 clusters was implemented. The test is referred to as a **Bootstrap likelihood ratio test (BLRT)** (McLachlan & Peel, 2000). "*Replication samples are generated from the probability distribution defined by the maximum likelihood estimates under H_0 (j-1-model)*" (Vermunt & Magidson, p. 54, 2005). The p-value reflects the proportion of

samples with a larger BLRT statistic than the original one. P-value less than 0.05 indicate a benefit of additional cluster for the model fit (McLachlan & Peel, 2000).

Choosing the optimal model: practical evaluation

It has been argued that in addition to judging goodness of fit, choosing the optimal model should encompass practical assessment of the data (Vermunt & Magidson, 2002, Magidson & Vermunt 2004, Peng, 2011). Overall, the criteria for pragmatic evaluation as summarised by Peng (2011) were followed, including:

- The optimal model is judged as optimal if had a large (near to 1) average posterior probability across clusters (e.g. 0.80, 0.10, 0.10), with no comparable average posterior probabilities for two or more clusters (e.g. 0.30, 0.30, 0.40).
- Clusters in the optimal model are expected to have distinct characteristics.
- There is no consensus on the minimum cluster size. The adequate size of a cluster depends on the aims and the sample size (Croudace et al., 2003, Clark et al., 2006, Nylund et al., 2007). In depression symptoms literature can be found examples of cluster size as small as 1.3% (Olino et al., 2010). Peng (2011) also adopted 1.0% minimum prevalence of class membership. The minimum cluster size for a dataset can be practically evaluated by observing of changes in models' stability (Peng, 2011).

Latent GOLD statistical software

For the purpose of model estimation, Latent GOLD relies on the mean of the maximum likelihood (ML) estimator of item conditional probabilities (Vermunt & Magidson, 2005). Latent GOLD generates the expectation-maximisation (EM) algorithm with random start values. EM is argued to be stable, and thus, a good estimator for

obtaining a converged solution (Vermunt & Magidson, 2005), defined as models having no difficulties in optimising the likelihood function across random start values (Peng, 2011). One of the possible problems with this method is the possibility that a converged solution might not be a global optimal solution (Vermut & Magidson, 2005), which is indicated by the number of times the largest log-likelihood is replicated (Peng, 2011). There are inconsistencies in numbers of replications indicating a local optimal solution (Múthen & Múthen, 2009, Hipp & Bauer, 2006). One recommended method of avoiding this problem is the use of multiple sets of starting values (Vermut & Magidson, 2005), with the disagreement of the optimal number required (Múthen & Múthen, 2005, Hipp & Bauer, 2006). Based on previous investigations (e.g. Kreuter & Múthen, 2008, Croudace et al, 2003) Peng (2011) has recommended 1000 starting values in running through 100 iterations.

5.3 BACKGROUND

5.3.1 Depression and anxiety symptoms over time: previous research in older adults and tertiary care patients with musculoskeletal problems

No studies investigating symptoms change over time for anxiety and depression symptoms in older people with musculoskeletal pain could be identified. However, potentially informative evidence was identified from studies that used predominately variable-centred methods to investigate the course of depressive and anxiety symptoms in older adults and adults with pain in the community. Also of use are the results of a person-centred analysis of the course of depressive and anxiety symptoms in tertiary care patients with chronic pain (including musculoskeletal origin) and rheumatoid arthritis.

There is robust evidence to suggest that the course of depression

symptoms in depressed older in the community is heterogeneous, and that symptoms often persist (Cole, 1999, Licht-Strunk et al., 2007). A systematic review found that in adults aged 60-65 years with depressive disorders, 32.7% had clinically significant symptoms of depression at 24 months (Cole et al., 1999). Cole et al. (1999) conclude that depression in population and primary care older adults has a poor prognosis and is likely to be chronic and/or, relapsing. Estimated rates of depression at follow-up in older adults who were depressed at baseline are high, for example: 63.2% at one year (Prince et al., 1998), 61.2% at two years (Harris et al., 2006) and 50.4% (Beekman et al., 2001) and 51.7% (Schoevers et al., 2003) at three years. Licht-Strunk et al. (2009a, 2009b) have demonstrated that the recovery rates of major depressive disorder in primary care older adults are low, but that they do increase with time; from 20.5% at 6 months to 35% at 1 year, 60% within 2 years and 68% in 3 years. The authors estimated that the median duration of a major depressive episode in older adults was 18 months (Licht-Strunk et al., 2009b). In a separate set of analyses Licht-Stunk investigated trajectories of depression symptoms over 3 years, in 296 older general practice attendees with major depressive disorder or sub-threshold depression assessed through at least four interviews (Licht-Strunk, 2008). The 4-cluster solution was chosen with 42% of patients having no depression symptoms from 6 months onwards (so-called *Recovery cluster*). The remaining three clusters include people who continued experiencing symptoms of depression for three years (Clusters: *Persistent mild symptoms* (35%), *Chronic symptoms* (18%), *Chronic severe symptoms* (5%)) (Licht-Strunk, 2008).

Anxiety data in older adults is scarce, but one large study indicates that anxiety symptoms might also have heterogeneous outcomes and a tendency to persist over time (De Beurs et al., 2000, Schuurmans et al., 2005). Overall, 26% of participants

had HADS defined elevated anxiety symptoms at baseline. Of these 17.9% were found to be chronically anxious over a 3-year period. Of those without anxiety symptoms at baseline, 10.8% were found to become anxious by a 3-year follow-up (De Beurs et al., 2000). Based on the same sample, Schuurmans et al. (2005) have estimated that 23% of the participants with anxiety disorders at baseline had persistent anxiety at a 6-year follow-up, with 47% having sub-syndromal symptoms and 31% in full recovery.

It is unclear if having OA or joint pain increases the risk of an adverse course of depressive and anxiety symptoms. Chronic physical illness (including OA) did not predict poor anxiety outcome in a large sample of older adults (De Beurs et al., 2000, Schuurmans et al., 2005). Cole et al.'s (1999) review concluded an inconsistent impact of physical illness on the course of depression. Supporting evidence is provided by a large study of adults with depressive and/or anxiety disorders and coexisting pain, recruited from community, general practice, and secondary mental health care settings (Gerrits et al., 2012). Overall, 61.5% of people with anxiety or depressive disorders at baseline still suffered from clinically significant symptoms depressive or anxiety at a 2-year follow-up (Gerrits et al., 2012). In addition, 43.5% had a chronic course with at least mild depression symptoms over the entire follow-up period. Pain location, high number of pain sites, ≥90 days of pain, using pain medication daily and a higher CPG score were significant predictors of depression symptoms at the 2-year follow-up (Gerrits et al., 2012). Pain for 90 days and a higher CPG were associated with having a chronic course, but after adjustment for the severity of depression/anxiety none of these factors were associated with 'chronic course' of depression symptoms. However, a large study of adults with early rheumatoid arthritis, recruited from rheumatology clinics, found that depressive or

anxiety symptoms can persist for over 10 years (Norton et al., 2011). A latent general mixture modelling identified four distinct trajectories of 'psychological distress', i.e. HADS defined depressive and/or anxiety symptoms, over 10 years including: low-stable HADS score (68%), high-stable HADS score (12%), high-decreasing HADS score (9%) and low-increasing HADS score (11%) (Norton et al., 2011). People with high-stable HADS symptoms had poor pain profile at baseline (i.e. more tender joints, higher pain intensity, early morning stiffness and higher disability) (Norton et al., 2011), but correlations between variables and their relative importance remains unclear as regression analyses were not performed.

5.3.2 The importance of examining symptom change over time

The clinical importance of examining symptom changes over time has been acknowledged in the field of mental health research (Solomon et al., 1997). For example, Lucassen et al. (2008) advocate a stepped model for the diagnosis of depression in primary care, including problem formulation by a patient, supporting the patient in problem solving, followed by a period of 'active monitoring' and re-assessment before deciding on the medical diagnosis and treatment. The understanding of the development of anxiety and depression over time and the associated factors is important for clinical decision making (NICE, 2009a). Clinicians have been found to experience difficulties in distinguishing between clinically significant cases and psychological problems of a transient nature (Barley et al., 2011, van Rijswijk et al., 2009). Consequently, it can help to estimate how many patients that due to the severity and persistence of their symptoms, might require medical intervention (Pedersen et al. 2008). A further important aspect of investigating heterogeneity of trajectories of symptoms is the value of this information for the patient, including increased

awareness of prognosis, i.e. how it can influence their life and what interventions might be received (Dunn et al., 2006).

5.4 RATIONALE OF THE STUDY

Clinical guidelines recommend that the management of depression symptoms should include consideration for symptom trajectories (NICE, 2009a), yet to date limited information is available to support clinicians managing patients with OA in the community. Qualitative data on primary care practitioners' views on the management of anxiety and depression in primary care (van Rijswijk et al., 2009) suggests that a limited understanding of the natural history of progress of depressive and anxiety symptoms acts as a barrier to their effective recognition and management. The analysis presented in this chapter will investigate the course of anxiety and depression symptoms in older people with musculoskeletal pain in primary care.

5.5 AIM AND OBJECTIVES

The **overall aim** of this study is to advance understanding of the persistence of depressive and anxiety symptoms in older primary care patients with OA.

Specific objectives are:

- To describe changes in the rate of HADS defined depressive and/or anxiety symptoms in older patients presenting to general practice with musculoskeletal pain
- To identify discrete 12-month post-consultation trajectories of symptoms of anxiety and depression in older patients presenting to general practice with musculoskeletal pain

- To explore patterns of coexisting trajectories of symptoms of anxiety and depression

5.6 METHOD

5.6.1 Sample

For consistency, descriptions of changes in the rate of symptom anxiety and depression symptoms reporting were based on the same samples used to choose the optimal models of depressive and anxiety trajectories. The choice of anxiety and depression models was based on participants who provided complete HADS anxiety and depression data (scores at all 4 time points). This way, individual HADS anxiety and depression values could be explored in each identified cluster. This would not be possible if a sample with missing values had been used, as the Latent GOLD software does not impute missing values.

5.6.2 Statistical analyses: sampling frame effects on the main outcomes and person-related characteristics

As a loss of the sample can raise concerns about the validity of the results (He, 2010) the possibility of selection bias was explored. The severity of anxiety and depression symptoms and baseline covariates included in subsequent chapters were described (using PASW version 18.0) for all baseline participants who consented to follow-up (n=443) and compared with individuals who provided complete HADS anxiety and depression data. The median was used for descriptive purposes of three baseline variables, due to non-normal distribution of scores (see Appendix D.1 on pages 374-376 for distribution of scores).

The effect of participants with complete data on the main outcome over time was further explored. The frequency of depressive and anxiety symptom severity were described for people who provided complete anxiety and depression data, and then compared against all participants (i.e. HADS anxiety and depression scores at 0-4 time points). To avoid multiple testing, statistical significance (established at the level of $p<0.05$) of differences between selected frequencies, was verified for the most apparent discrepancies (followed by testing all differences only if the most apparent differences were found to be statistically significant). For this purpose the Chi^2 test or the Fisher's exact test (when expected cell frequencies were less than 5) was used. Data analyses were performed using PASW version 18.0.

5.6.3 Statistical analyses: analyses to describe changes in the rate of HADS defined depressive and anxiety symptoms

The course of the rate of depressive and anxiety symptoms was explored in participants with complete depressive and anxiety data by describing changes in the frequencies of symptom reporting. Data was grouped by severity of symptoms at baseline and proportions of different severity of depression and anxiety over time 0-4 were described. Analyses were conducted in PASW version 18.0.

5.6.4 Statistical analyses: discrete trajectories of depressive and/or anxiety symptoms

One of the assumptions of LCGA is a multivariate normal distribution of parameters describing individual level trajectories (Titterington et al., 1985). As this assumption was not met in the investigated sample (Table 4.3. on page 146) a

non-distribution based method of model estimation was chosen – as suggested by Bauer and Curran (2003). It involved dichotomising data (HADS scores 0-7 vs. ≥8). A practical reason for this decision was a relatively small sample size, examples of previous research (e.g. Licht-Strunk, 2008) allowed expecting that analyses based on continuous data were likely to result in a number of small clusters (increasing the risk of model instability). This decision is supported with clinical reasons, at least for depression, any elevated symptoms are of clinical relevance to UK primary care (NICE, 2009b).

Individual patterns of binary depression and anxiety outcomes at 4 time points were generated (using PASW version 18.0). LCGA models were then estimated using Latent GOLD, using the mean of the maximum likelihood (ML) estimator of item conditional probabilities (Vermunt & Magidson, 2005). To generate the expectation-maximization (EM) algorithm, 1000 starting values running through 100 iterations were used for LCGA models (Croudace et al, 2003, Kreuter & Muthén, 2008, Peng, 2011). Three indicators of goodness of fit were generated with the Latent GOLD including: log-likelihood (LL)**,** Bayesian Information Criterion (BIC) and Bootstrap likelihood ratio test (BLRT) (with statistical significance established at the level of $p<0.05$). The posterior membership probabilities derived from Baye's theorem were then used to classify cluster membership (Peng, p. 41, 2011). Average posterior probabilities were then calculated for each cluster using PASW version 18.0. Sensitivity analyses, involving multiple repetitions of the above procedure, revealed that clusters of less than 5.0 % prevalence tended to be unstable, and so a minimum of ≥5.0% of cluster prevalence was decided for evolution of practical utility of a LCGA model. To investigate associations between the selected models

of depressive and anxiety trajectories, prevalence rates of the depression trajectories nested in the anxiety model were estimated.

5.7 RESULTS

5.7.1 Sample size

Of the 443 baseline consenters for follow up, 293 (66% of the consenters) and 298 (67%) participants had HADS anxiety scores and HADS depressive scores respectively, available at all four time points.

5.7.2 Sampling frame effects on the main outcomes and the covariates

The possibility of bias resulting from omitting participants was explored. Outcome variables at baseline, and baseline covariates, were compared across the original sample of consenters for follow-up (n=443), samples used for choosing the optimal anxiety and depression models (n=298 and n=293 respectively). Table 5.2 overleaf shows that differences were marginal indicating that substantial selection bias was unlikely.

Table 5.2 Baseline characteristics for consenters for follow-up and selected samples used in chapter six.

Factors	Consented for follow-up n=443 n (%)	Anxiety symptoms Complete data over time n=293 n (%)	Depression symptoms Complete data over time n=298 n (%)
Main outcome:			
HADS-D			
None (0 - 7)	315 (72.1)	215 (73.6)	219 (73.5)
Mild (8 - 10)	69 (15.8)	47 (16.1)	47 (15.8)
Moderate (11 - 14)	43 (9.8)	27 (9.2)	29 (9.7)
Severe (15 - 21)	10 (2.3)	3 (1.0)	3 (1.0)
HADS-A			
None (0 - 7)	234 (53.8)	151 (51.5)	149 (50.5)
Mild (8 - 10)	106 (24.4)	80 (27.3)	79 (26.8)
Moderate (11 - 14)	77 (17.7)	51(17.4)	19 (19.0)
Severe (15 - 21)	18 (4.1)	11 (3.8)	11 (3.7)
Missing	8	-	3
Pain characteristic: median (IQR)			
Number of pain sites (0-44)	7 (9)	7 (9)	7 (9)
Interference with daily activities (0-10)	6 (5)	5(4)	5 (10)
Missing	18	9	11
Interference with work (0-10)	5 (5)	5 (4)	5 (5)
Missing	16	7	9
Interference with social activities (0-10)	6 (5)	5 (5)	5 (5)
Missing	15	6	8

Table 5.2 cont. Baseline characteristics for consenters for follow-up and selected samples used in chapter six.

Factors	Consented for follow-up n=443 n (%)	Anxiety symptoms Complete data over time n=293 n (%)	Depression symptoms Complete data over time n=298 n (%)
Socio-ecological:			
Living alone:			
No	361 (81.5)	244 (83.3)	244 (81.9)
Yes	80 (18.1)	49 (16.7)	54 (18.1)
Missing	2	-	-
Marital status:			
Married	317(72.2)	217 (74.6)	217 (73.3)
Separated	6 (1.4)	4 (1.4)	4 (1.4)
Divorced	25 (5.7)	11(3.8)	12 (4.1)
Widowed	65 (14.7)	41 (14.1)	45 (15.2)
Cohabiting	16 (3.6)	11 (3.8)	11 (3.7)
Single	10 (2.3)	7 (2.4)	7 (3.7)
Missing	4	2	2
Instrumental support:			
Yes	348 (78.9)	225 (76.8)	229 (76.8)
No	35 (7.9)	25 (8.5)	27 (9.1)
No need	58 (13.2)	43 (14.7)	42 (14.1)
Missing	2	-	-
Emotional support:			
Yes	385 (87.9)	255 (87.6)	260 (87.8)
No	21 (4.7)	11 (3.8)	11 (3.7)
No need	32 (7.3)	25 (8.6)	25 (8.4)
Missing	5	2	2
Socio-economic status:			
Managerial/professional	126 (28.4)	86 (29.4)	88 (29.5)
Intermediate	92 (20.8)	72 (24.6)	73 (24.5)
Routine/manual	162 (36.6)	103 (35.2)	103 (34.6)
Other	63 (14.2)	32(10.9)	34 (11.4)

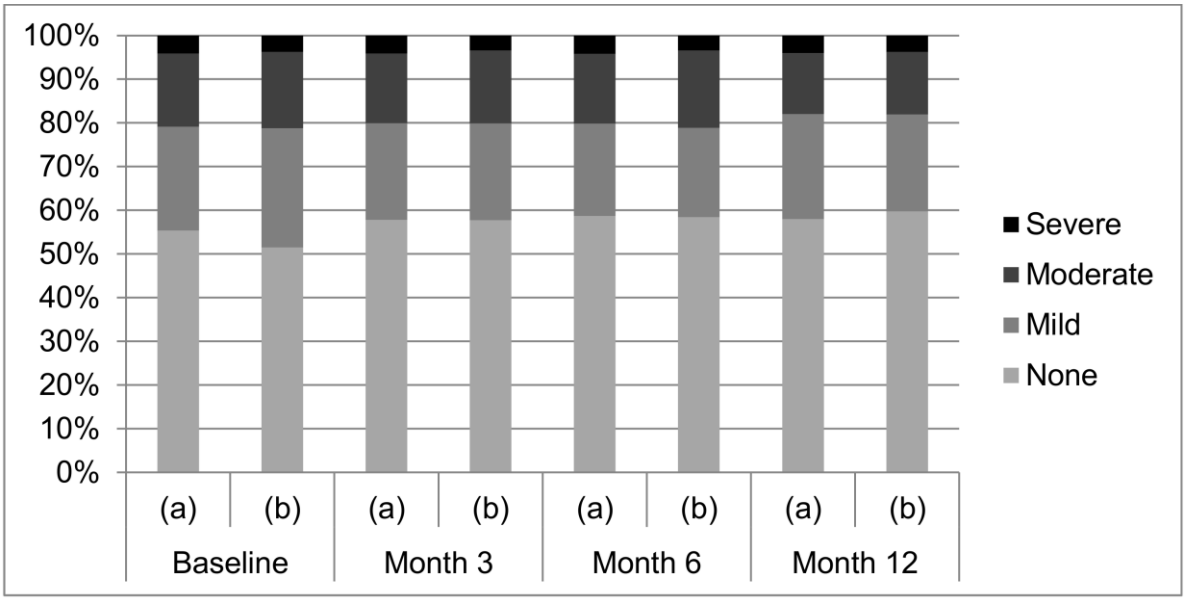
Table 5.2 cont. Baseline characteristics for consenters for follow-up and selected samples used in chapter six.

Factors	Consented for follow-up n=443 n (%)	Anxiety symptoms Complete data over time n=293 n (%)	Depression symptoms Complete data over time n=298 n (%)
Demographic characteristics:			
Age			
50-59 years	162 (36.6)	96 (32.8)	96 (32.2)
60-69 years	146 (33.0)	110 (37.5)	112 (37.6)
70-79 years	98 (22.1)	70 (23.9)	69 (23.2)
80+ years	37 (8.3)	17 (5.8)	21 (7.0)
Gender			
Female gender	271 (61.2)	177(60.4)	182 (61.1)
Coping^:			
Catastrophising (0-6)			
<(4) highest tertile	282 (69.1)	194 (70.8)	197 (71.1)
Missing	35	19	21
Self-statements (0-6)			
<(4.5) highest tertile	227 (55.5)	152 (54.5)	156 (55.1)
Missing	34	14	15
Ignoring pain (0-6)			
<(4) highest tertile	288 (71.8)	199 (73.7)	202 (73.7)
Missing	42	23	24
Increased behavioural activities (0-6)			
<(5) highest tertile	192 (47.2)	129 (47.3)	133 (47.8)
Missing	36	20	20

Note: ^ - highest tertiles for coping are based on n=502; HADS - Hospital Anxiety and Depression Scale; IQR- inter quartile range.

The severity of anxiety symptoms at each time point for the 293 participants who provided complete anxiety data (i.e. HADS anxiety score at all 4 time points) were described and then compared against all participants (i.e. HADS anxiety score at 0-4 time points). Results are reported in Figure 5.1 and the associated table. The most apparent difference was found for the baseline proportions of people without anxiety symptoms (54.2% vs. 51.5%), yet this difference was statistically not significant (chi^2= 1.04, p=0.308). This indicates that using complete anxiety data was unlikely to result in substantial selection bias.

Figure 5.1 Prevalence of none, mild, moderate and severe anxiety symptoms over time, for all patients (a) and with complete anxiety data at 4 time points (b).

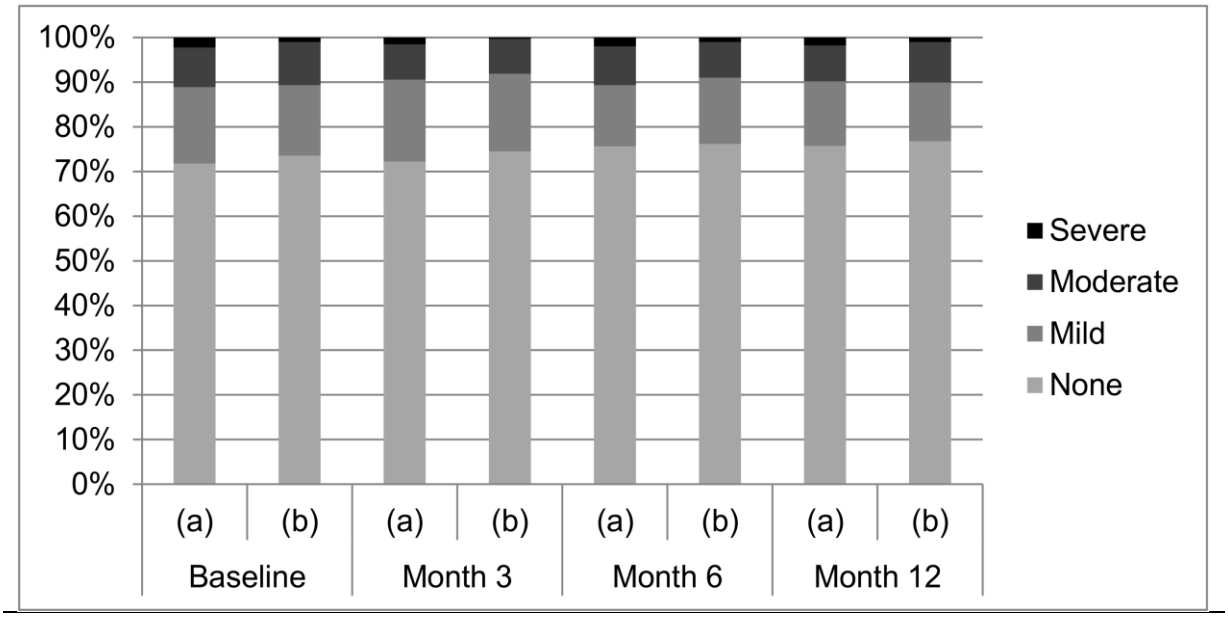


	Severity of anxiety symptoms (HADS-A score)				
	None (0-7) n (%)	**Mild (8-10) n (%)**	**Moderate (11-14) n (%)**	**Severe(15-21) n (%)**	**Total sample**
(a) All participants^					
Baseline	272 (54.2)	117 (23.3)	83 (16.5)	20 (4.0)	502
Month 3	225 (56.5)	86 (21.6)	62 (15.6)	16 (4.0)	398
Month 6	210 (56.8)	76 (20.5)	57 (15.4)	15 (4.1)	370
Month 12	188 (57.1)	78 (23.7)	45 (13.7)	13 (4.0)	329
(b) Participants with complete anxiety data					
Baseline	151 (51.5)	80 (27.3)	51 (17.4)	11 (3.8)	293
Month 3	169 (57.7)	65 (22.2)	49 (16.7)	10 (3.4)	293
Month 6	171 (58.4)	60 (20.5)	52 (17.7)	10 (3.4)	293
Month 12	175 (59.7)	65 (22.2)	42 (14.3)	11 (3.8)	293

Note: ^- numbers do not add up to 100% due to missing data.

The severity of depression symptoms at each time point in the 298 participants who provided complete depression data (i.e. HADS-D score at all 4 time points) were described and compared against all participants (i.e. HADS-D score at 0-4 time points) (Figure 5.2 and the assistant table). The most apparent differences were in the proportions of severely depressed patients between the two groups, with Fisher's exact tests indicating statistically non-significant differences at baseline (p=0.272), 3 (p=0.250), 6 (p=0.360) and 12 (p=0.508) months.

Figure 5.2 Prevalence of none, mild, moderate and severe depression symptoms over time, for all patients (a) and with complete depression data all time points (b).



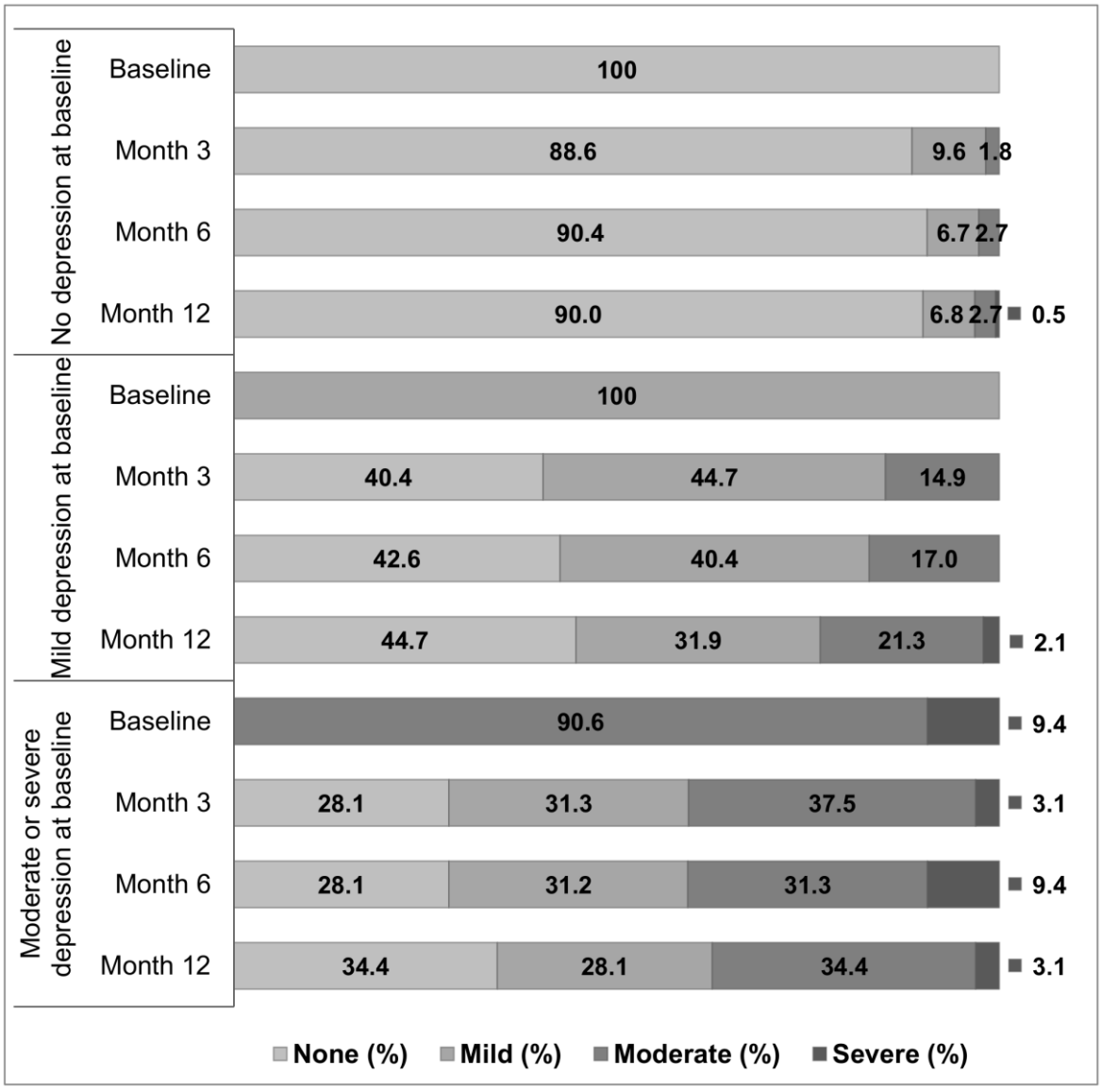
	Severity of depression symptoms (HADS-D score)				
	None (0-7) n (%)	**Mild (8-10) n (%)**	**Moderate (11-14) n (%)**	**Severe (15-21) n (%)**	**Total sample**
(a) All participants^					
Baseline	356 (70.9)	85 (16.9)	44 (8.8)	11(2.2)	502
Month 3	283 (71.1)	72 (18.1)	31 (7.8)	6 (1.5)	398
Month 6	270 (73.0)	49 (13.2)	31 (8.4)	7 (1.9)	370
Month 12	246 (74.8)	47 (14.3)	26 (7.9)	6 (1.8)	329
(b) Participants with complete depression data					
Baseline	219 (73.5)	47 (15.8)	29 (9.7)	3 (1.0)	298
Month 3	222 (74.5)	52 (17.4)	23 (7.8)	1 (0.3)	298
Month 6	227 (76.2)	44 (14.8)	24 (8.0)	3 (1.0)	298
Month 12	229 (76.8)	39 (13.1)	27 (9.1)	3 (1.0)	298

Note: ^- numbers do not add up to 100% due to missing data.

5.7.3 Changes in the rate of depression symptoms over time

Changes in the frequencies of none, mild, moderate and severe symptoms of depression were calculated for patients with complete depression data (n=298) at four time points. This was done separately for those without depression at baseline (n=219), those with mild symptoms at baseline (n=47) and those with moderate or severe symptoms (n=32). Results can be found displayed in Figure 5.3.

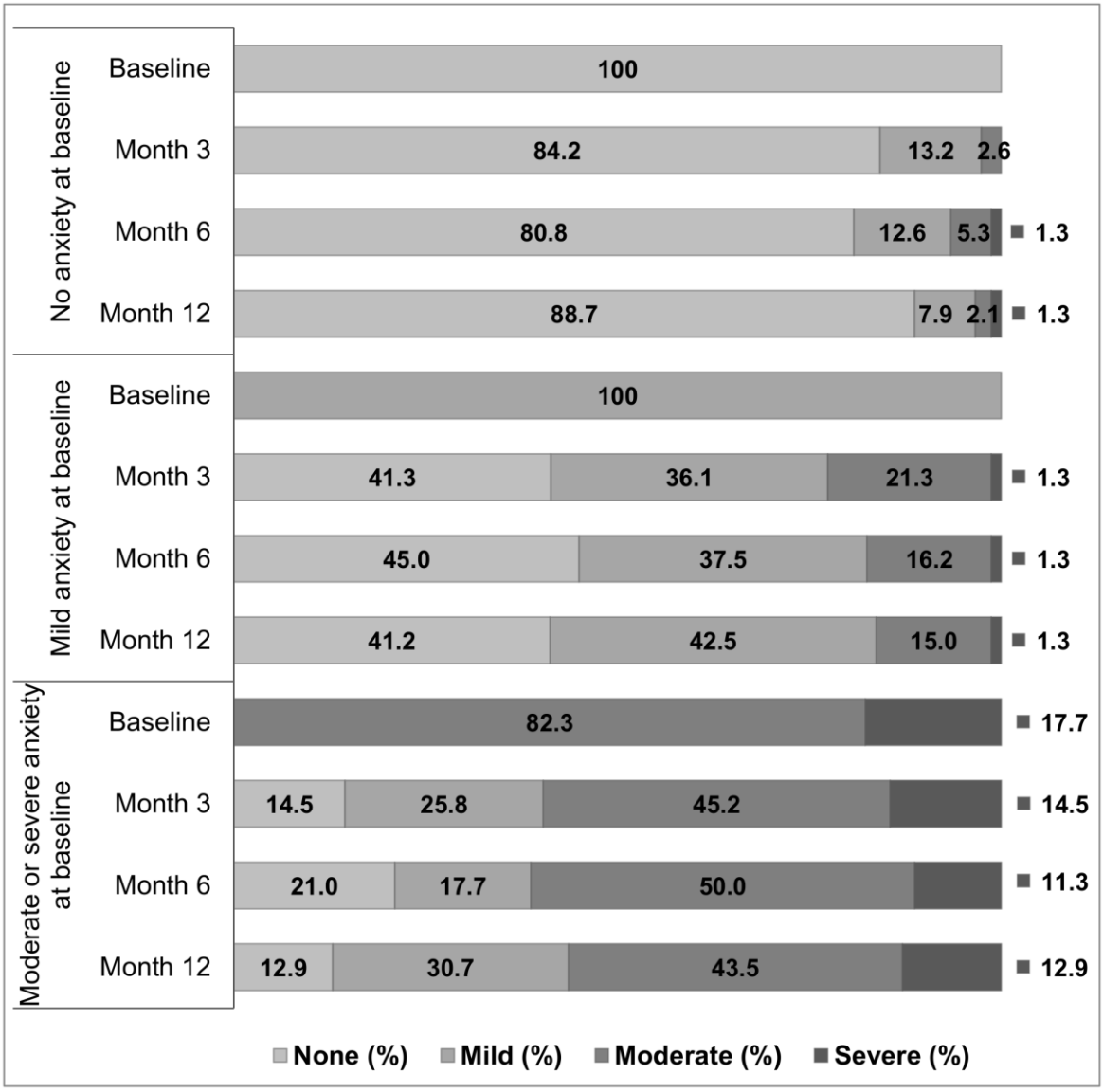
Figure 5.3 Frequencies of symptoms severity at 0-4 time points for patients without depression at baseline.



5.7.4 Changes in the rate of anxiety symptoms over time

Patients with complete anxiety data at four time points (n=293) were split into three groups by severity of symptoms at baseline, including none (n=151), mild (n=80) and moderate or severe symptoms (n=62). Frequencies of none, mild, moderate and severe symptoms were then calculated for each group at 3, 6 and 12 months. Results can be found displayed in Figure 5.4.

Figure 5.4 Frequencies of symptom severity at 0-4 time points for patients without anxiety at baseline.



5.7.5 Trajectories of depression symptoms

Description of the identified individual patterns of depression symptoms

Exemplars of the observed individual patterns of HADS-anxiety or -depressive scores are displayed in Appendix D.2 (on page 376). Patterns of binary depression data were examined for the 298 participants with complete data (Table 5.3), with 16 patterns revealed, including 12 patterns that occurred ≥1.0% of participants. The most prevalent was a pattern of no depression at four time points (58.4%), followed by the pattern of depression at all four time points (12.4%). Some other common patterns such as '1000' (6.0%) and '0100' (3.6%) represent participants who at baseline and 3 months reported depression symptoms and the symptoms resolved by the 12-month follow-up.

Table 5.3 The observed patterns of numbers of participants with (1) and without (0) elevated symptoms of depression across the four time points (n=298).

Pattern*	N	%
0000	174	58.4
1111	37	12.4
1000	18	6.0
0100	12	4.0
0001	10	3.6
0010	7	2.3
0111	7	2.3
1110	7	2.3
1100	6	2.0
1011	5	1.7
1001	4	1.3
0110	4	1.3
0011	3	1.0
0101	2	0.8
1010	1	0.3
1101	1	0.3
Total	298	100%

Note: *- A pattern is a combination of binary grouping of HADS-D scores (0= scores <8, 1= scores ≥ 8) with an order of baseline, 3 month, 6 month, 12 month; These patterns comprise 67.3% of baseline respondents who consented for follow-up.

LCGA depression models

1,2...5-cluster LCGA were fitted to binary depression data. The results are presented in Table 5.4. BIC was the lowest for the 2-cluster model solution. The bootstrap likelihood ratio test showed statistical significance between the 1 cluster and 2-cluster model (p<0.0001). This indicates the benefit of adding a cluster to a one cluster solution. Statistically non-significant effects were observed by comparing the 2- and 3-cluster models, 3- and 4-cluster models and 4-and 5-cluster models.

Table 5.4 Optimal number of clusters: goodness of fit statistics for LCGA models using 4 time points and complete binary depression data (n=298).

Cluster	LL	Par.	BIC(LL)	BLRT	(p-value)
1	-666.4341	3	1349.9594	NA	
2	-492.9665	7	1025.8126	346.9352	(<0.0001)
3	-488.7824	11	1040.2329	8.3680	(0.0680)
4	-486.3759	15	1058.2081	4.8132	(0.1880)
5	-486.2550	19	1080.7547	0.2418	(0.4400)

Note: LL- Log-likelihood; Par.- no. of parameters; BIC- Bayesian Information Criterion; BLRT=bootstrap parametric likelihood ratio test: k- cluster model vs. (k-1) cluster (BLRT values are not available for single group models (NA)).

Figure 5.5 Log-Likelihood (LL) for 1-5 clusters depression LCGA models.

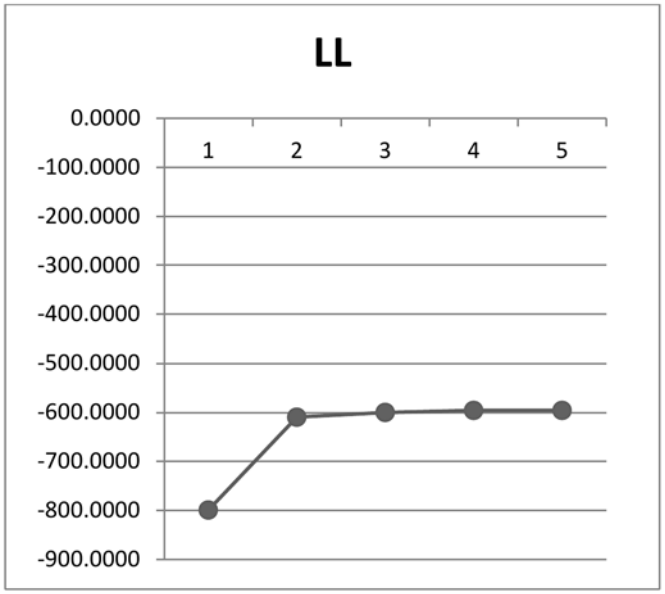


Figure 5.5 shows that the log-likelihood (LL) flattened after the 2-cluster model, suggesting that the 2-cluster solution was optimal. Overall, assessment statistics for the LCGA model of binary depression outcome among eligible participants consistently suggest that the 2-cluster depression model is optimal.

Following the results of the goodness of fit index, the 2-cluster depression model was explored by assessing the proportion of individuals in each cluster and the distinctiveness of clusters, using average posterior probability. The smallest cluster consisted of more than 5.0% of the total sample, and so had an adequate sample size. Both clusters had high average posterior probabilities (0.9779, 0.9579) for belonging to the assigned cluster and average posterior probability of belonging to the other cluster were low (Table 5.5 overleaf). This indicates a good separation of cluster 1 from cluster 2. It was therefore decided that the 2-cluster model was the optimal LCGA model.

Table 5.5 Average assignment probabilities based on maximum posterior probability for 2 clusters LCGA model for complete depression data.

Assigned Cluster	n	%	Average posterior probabilities for each cluster	
			1	2
1	232	77.8	**0.9779**	0.0221
2	66	22.2	0.0421	**0.9579**

The proportions of observed patterns of depression symptoms in each cluster are listed in Table 5.6 overleaf. Cluster one (Figure 5.6 on page 183), referred to as the *no depression symptom trajectory*, comprised the majority of the sample (77.8%) and predominantly included people without depression symptoms or with transient depression symptoms over time. Cluster two (Figure 5.7 on page 184), referred to here as *the persistent depression symptom trajectory*, included patients with persistent depression symptoms (22.2%). An individual with a reported '1' or '0' at three time points was likely to be classified to the same cluster as a person with the same outcome at all 4 time points. When '1' or '0' were recorded twice, a classification was associated with a sequence of occurrence, so that, when depression symptoms emerged at 3, 6 or 12 months a person was likely to be classified as persistently depressed. It can be argued that those patterns can form a separate, small group, but this was not subsequently supported by the goodness of fit statistics.

Table 5.6 Patterns* observed in clusters 1 and 2 of the 2-cluster LCGA depression model.

Cluster 1 n=222 (no depression)			Cluster 2 n=66 (persistent depression symptoms)		
Pattern	n	(%)	Pattern	n	(%)
0000	174	(75.0)	1111	37	(56.0)
1000	18	(7.8)	1110	7	(10.6)
0100	12	(5.2)	0111	7	(10.6)
0001	10	(4.3)	1011	5	(7.6)
0010	7	(3.0)	0110	4	(6.1)
1100	6	(2.6)	0011	3	(4.5)
1001	4	(1.7)	0101	2	(3.0)
1010	1	(0.4)	1101	1	(1.5)

Note: *- A pattern is a combination of binary grouping of HADS-A scores with an order of baseline, 3, 6 and 12 month follow-ups.

Figure 5.6 overleaf shows the proportion of patients in the *no depression symptom trajectory*, with different classifications of symptom severity across four time points being displayed. Amongst the 232 patients in this cluster, twenty nine had ‘mild or worse’ symptoms at baseline. This number has gradually decreased from 18 to 8 cases at 6 months, followed by remission of 6 individuals into mild symptom severity at 12 months (see Figure D.3.1 in Appendix D.3 on page 379 for observed trajectories in the *no depression symptom trajectory*).

Figure 5.6 2-cluster LCGA depression model: *no depression symptom trajectory* (n=232).

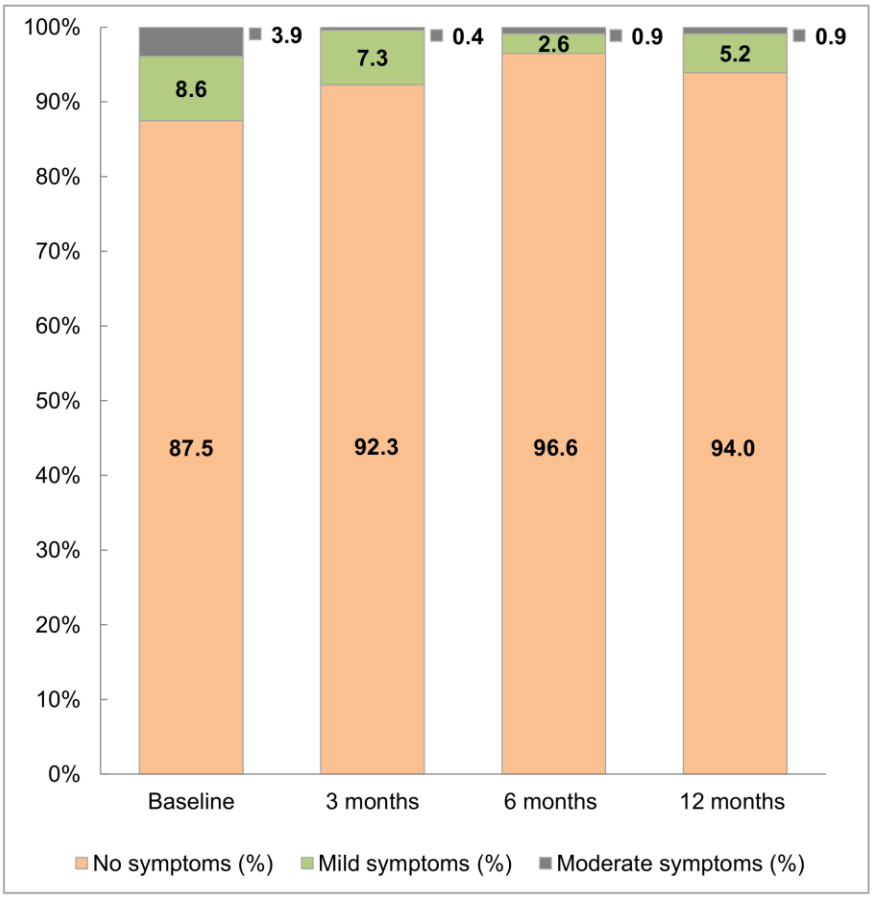
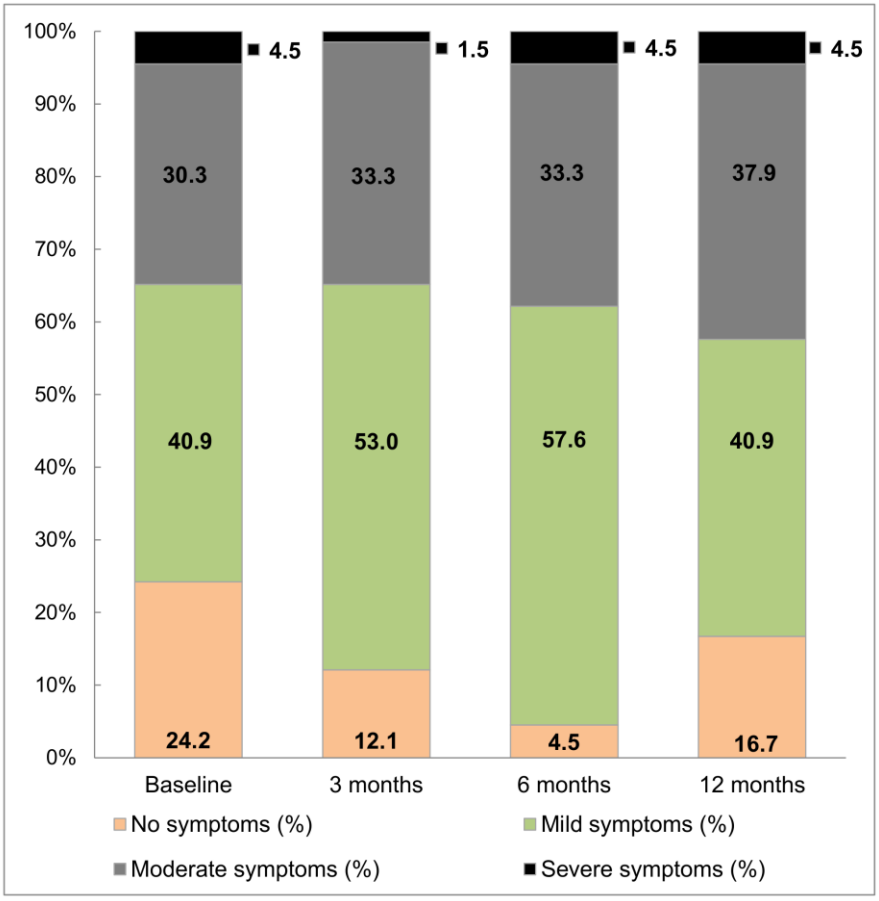


Figure 5.7 overleaf shows the proportion of patients in the *persistent depression symptom trajectory*, with different classifications of symptom severity across four time point. No depression symptoms were reported by 16, 8, 3 and 11 individuals at baseline, 3, 6 and 12 months respectively. Mild symptoms were reported by 27 (baseline), 35 (3 months), 38 (6 months) and 27 (12 months) patients. Moderate or severe symptoms were presented by 23, 23, 25 and 28 patients at baseline and three follow-ups respectively (see Figure D.3.2 in Appendix D.3 on page 379 for observed trajectories in the *persistent depression symptom trajectory*). Overall this cluster is characterised by increasing severity of symptoms from baseline to 6 months, with a decrease at 12 months. More

specifically, from 38 with mild symptoms at 6 months, 8 people showed a decrease of symptoms to no symptoms at 12 months and 2 developed into moderate symptoms at 12 months.

Figure 5.7 2-cluster LCGA depression model: *persistent depression symptom trajectory* (n=66).



5.7.6 Trajectories of anxiety symptoms

Description of the identified individual patterns of anxiety symptoms

Patterns of binary anxiety data were examined in the 293 participants with complete data, with 16 patterns revealed, of which 15 had frequency of occurrence equal or greater than 1.0% (Table 5.7 overleaf). The most prevalent pattern was of no anxiety at four time points (37.5%), followed by the pattern of persistent anxiety symptoms at all four time points (24.9%). Some other common patterns include

participants who initially reported symptoms 1000 (6.5%) and the symptoms resolved by 6 months and also participants with consistent anxiety at all but at 6 months (3.8%).

Table 5.7 Observed patterns of numbers of participants with (1) and without (0) elevated symptoms of anxiety across the four time points (n=293).

Pattern*	n	%
0000	110	37.5
1111	73	24.9
1000	19	6.5
1101	11	3.8
0010	10	3.4
1001	9	3.1
1100	9	3.1
0111	8	2.7
0100	8	2.7
0110	7	2.4
1011	7	2.4
1110	7	2.4
1010	6	2.0
0001	4	1.4
0011	4	1.4
0101	1	0.3
Total	293	100%

Note: *- A pattern is a combination of binary grouping of HADS-D scores (0= scores <8, 1= scores ≥8) with an order of baseline, 3 month, 6 month, 12 month;

These patterns comprise 66.1% of baseline respondents who consented for follow-up.

LCGA anxiety models

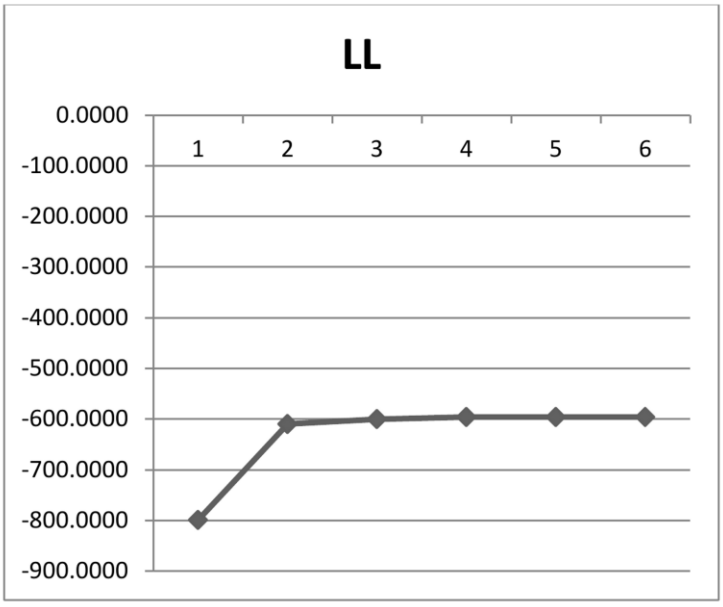
1-,2-,...6-cluster LCGA models were fitted to binary anxiety data and the results of assessment statistics are listed in Table 5.8. The log-likelihood values flattened out after the 2-cluster model (Figure 5.8). The same 2-cluster anxiety model was supported by BIC.

Table 5.8 Optimal number of clusters: assessment statistics for LCGA models using 4 time points and complete binary anxiety data (n=293).

Cluster	LL	Par.	BIC(LL)	BLRT	(p-value)
1	-799.1747	3	1615.3900	NA	
2	-609.7106	7	1259.1825	378.9282	(<0.0001)
3	-599.9943	11	1262.4704	19.4328	(<0.0001)
4	-595.9153	15	1277.0332	8.1579	(0.0260)
5	-595.8863	19	1299.6959	0.0580	(0.4800)
6	-595.8683	23	1322.3805	0.0361	(0.7500)

Note: LL- Log-likelihood, Par.- no. of parameters, BIC-Bayesian Information Criterion; BLRT= bootstrap parametric likelihood ratio test: k-cluster model vs. (k-1) cluster (BLRT values are not available for single group models (NA)).

Figure 5.8 Log-Likelihood (LL) for 1-5 clusters anxiety LCGA models.



The bootstrap likelihood test (BLRT) suggested a statistically significant difference between 1- and 2-cluster models, 2- and 3-cluster models, and also between 3- and 4- cluster solutions (Table 5.8 on the previous page). This indicates the benefit of adding a cluster to a 1- cluster, 2-cluster and 3-cluster solutions, but BLRT has modest significance for step up from 3 to 4. Therefore, the average posterior probabilities and the proportion of people within each cluster, across 2-, 3- and 4-cluster models were investigated.

The results of comparisons for the three clusters are presented in Table 5.9 overleaf, with average posterior probability marked in bold. In each model the smallest cluster was greater than 5.0% suggesting adequate cluster size. Each cluster in the 2-cluster anxiety model had a high average posterior probability (0.9247, 0.9641), which suggest good classification distinctiveness. Similarly, the 3-cluster model was characterised by high average posterior probability across all the clusters (0.8656, 0.9312, 0.8461). Clusters in the 4-cluster model had high posterior probability across three clusters (0.9528, 0.9048, 0.7847, 0.6392). Cluster-4 in the 4-cluster model had the lowest average posterior probability (0.6392), but average posterior probabilities for belonging to the other clusters were low (0.1721, 0.1260, 0.0628). Average posterior probabilities suggested that all three models were therefore considered acceptable. The results of the goodness of fit statistics together with average posterior probabilities, suggested that either the 3-cluster and 4-cluster anxiety models may be a more optimal solution than the 2-cluster anxiety model. Consequently, the decision was made to investigate characteristics of these two models.

Table 5.9 Average assignment probabilities based on maximum posterior probability for LCGA models for complete binary anxiety data (n=293).

2-cluster LCGA model

Assigned cluster	n	%	Average posterior probabilities for each cluster	
			1	2
1	159	54.3	**0.9641**	0.0359
2	134	45.7	0.0753	**0.9247**

3-cluster LCGA model

Assigned cluster	n	%	Average posterior probabilities for each cluster		
			1	2	3
1	119	40.6	**0.8656**	0.1342	0.0002
2	86	29.4	0.0272	**0.8461**	0.1267
3	88	30.0	0.0000	0.0688	**0.9312**

4-cluster LCGA model

Assigned cluster	n	%	Average posterior probabilities for each cluster			
			1	2	3	4
1	121	41.3	**0.9528**	0.0002	0.0381	0.0089
2	87	29.7	0.0002	**0.9048**	0.0948	0.0001
3	56	19.1	0.1196	0.0920	**0.7847**	0.0037
4	29	9.9	0.1721	0.1260	0.0628	0.**6392**

Table 5.10 overleaf shows patterns observed in clusters 1-3 in the 3-cluster solution. The most prevalent pattern in the 3-cluster model was cluster one (40.6%), which included individuals without elevated anxiety symptoms over time, referred to as the *no anxiety symptom trajectory* (Figure 5.9 on page 191). Cluster two (29.4%) included people of whom the majority experienced elevated symptoms at one or two time points, referred to as the *transient anxiety symptom trajectory* (Figure 5.10 on page 192). Comparable in size cluster three (30.0%) consisted of people with persistent anxiety symptoms (i.e. at three or more time points), hence called the *persistent anxiety symptom trajectory* (Figure 5.11 on page 193).

Table 5.10 Patterns* observed in clusters 1-3 of the 3-cluster LCGA anxiety model.

Cluster 1 n=119 (no symptoms)		Cluster 2 n=86 (transient anxiety symptoms)		Cluster 3 n=88 (persistent anxiety symptoms)	
Pattern	**n (%)**	**Pattern**	**n (%)**	**Pattern**	**n (%)**
0000	109 (91.6)	1000	19 (22.1)	1111	73 (83.0)
0010	10 (8.4)	1101	11 (12.8)	0111	8 (9.1)
		1001	10 (11.6)	1011	7 (8.0)
		1100	9 (10.5)		
		0100	8 (9.3)		
		0110	7 (8.1)		
		1110	7 (8.1)		
		1010	6 (7.0)		
		0001	4 (4.7)		
		0011	4 (4.7)		
		0101	1 (1.2)		

Note: * A pattern is a combination of binary grouping of HADS-A scores with an order of baseline, 3, 6 and 12 month follow-ups.

Patterns observed in clusters 1-4 in the 4-cluster solution are displayed in Table 5.10 (alternatively see Figure D.4.4 in Appendix D.4 on page 381). The three most prevalent clusters in the 4-cluster model were the 3 clusters nested in the 3-cluster model. The fourth cluster in the 4-cluster anxiety model, included 10 individuals previously allocated to cluster one in the 3-cluster model (pattern: 0010), 11 from cluster two in the 3-cluster model (7 with pattern: 0110; 4 with pattern: 0011) and 8 from cluster three in the 3-cluster model (pattern: 0111). The 4th cluster in the 4-cluster model appeared to have no distinct trajectory of anxiety symptoms, so the 3-cluster model was considered the optimal solution.

Table 5.11 Patterns* observed in clusters 1-4 of the 4-cluster LCGA anxiety model.

Cluster 1 n=121 (no symptoms)		Cluster 2 n=87 (persistent anxiety symptoms)		Cluster 3 n=56 (transient anxiety symptoms)		Cluster 4 n=29 (?)	
Pattern	**n (%)**	**Pattern**	**n (%)**	**Pattern**	**n (%)**	**Pattern**	**n (%)**
0000	109 (90.1)	1111	73 (84.0)	1000	19 (33.9)	0010	10 (34.5)
0100	8 (6.6)	1110	7 (8.0)	1101	11 (19.6)	0111	8 (27.6)
0001	4 (3.3)	1011	7 (8.0)	1001	10 (17.9)	0110	7 (24.1)
				1100	9 (16.1)	0011	4 (13.8)
				1010	6 (10.7)		
				0101	1 (1.8)		

Note: * - A pattern is a combination of binary grouping of HADS-A scores with an order of baseline, 3, 6 and 12 month follow-ups.

Figure 5.9 overleaf shows the proportions of patients with the *no anxiety symptom trajectory* and different classifications of symptoms over time. Anxiety symptoms were absent over time, except for 10 participants with anxiety symptoms at 6 months follow-up respectively. For individual patterns of HADS-anxiety scores over time see Figure D.5.1 in Appendix D.5 (on page 382).

Figure 5.9 3-cluster LCGA anxiety model: *no anxiety symptom trajectory* (n=119).

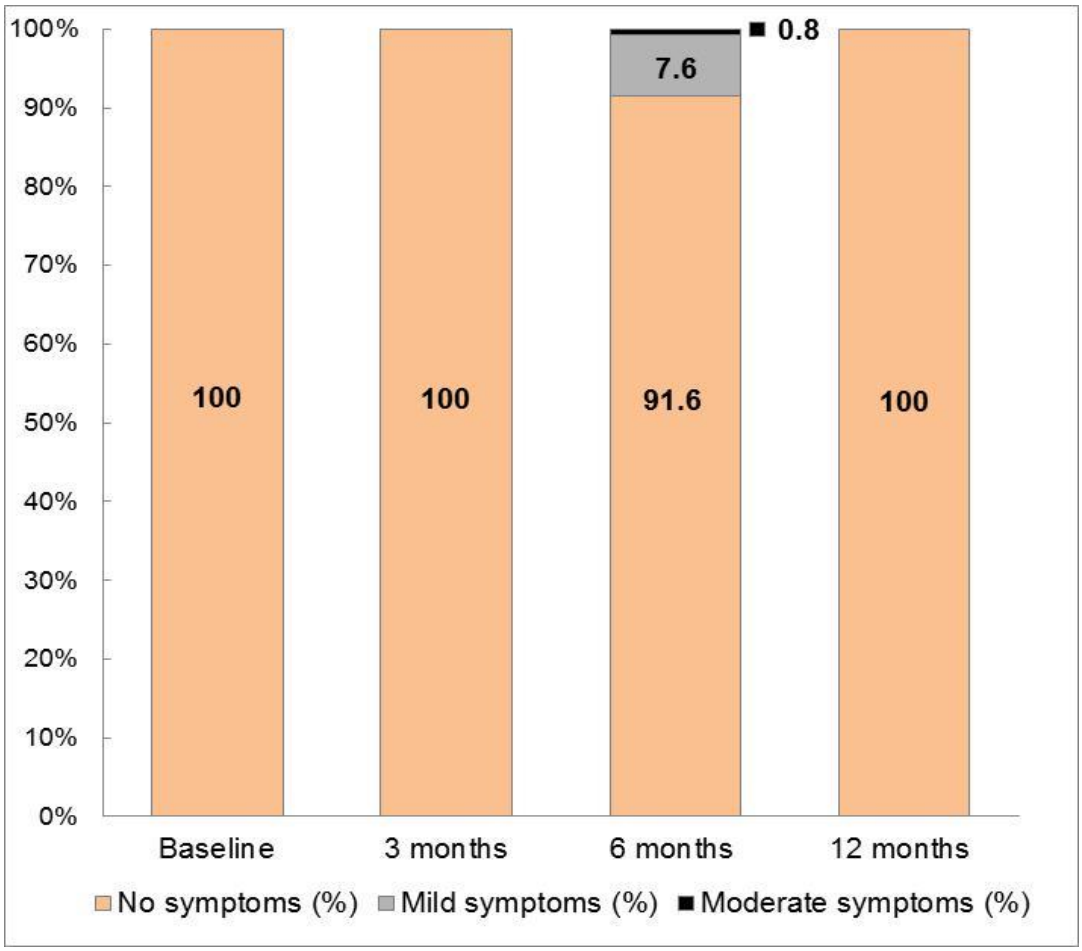


Figure 5.10 overleaf shows the proportions of patients with the *transient anxiety symptom trajectory* and different classifications of symptoms over time. Numbers of patients without anxiety symptoms increased from 26 individuals at baseline to 43 at 3 months and 62 at 6 months, this was followed by a decrease to 56 patients not anxious at 12 months. When elevated, symptoms of anxiety were predominantly mild (see Figure D.5.2 in Appendix D.5 on page 382 for individual patterns of HADS-anxiety scores over time).

Figure 5.10 3-cluster LCGA anxiety model: *transient anxiety symptom trajectory* (n=86).

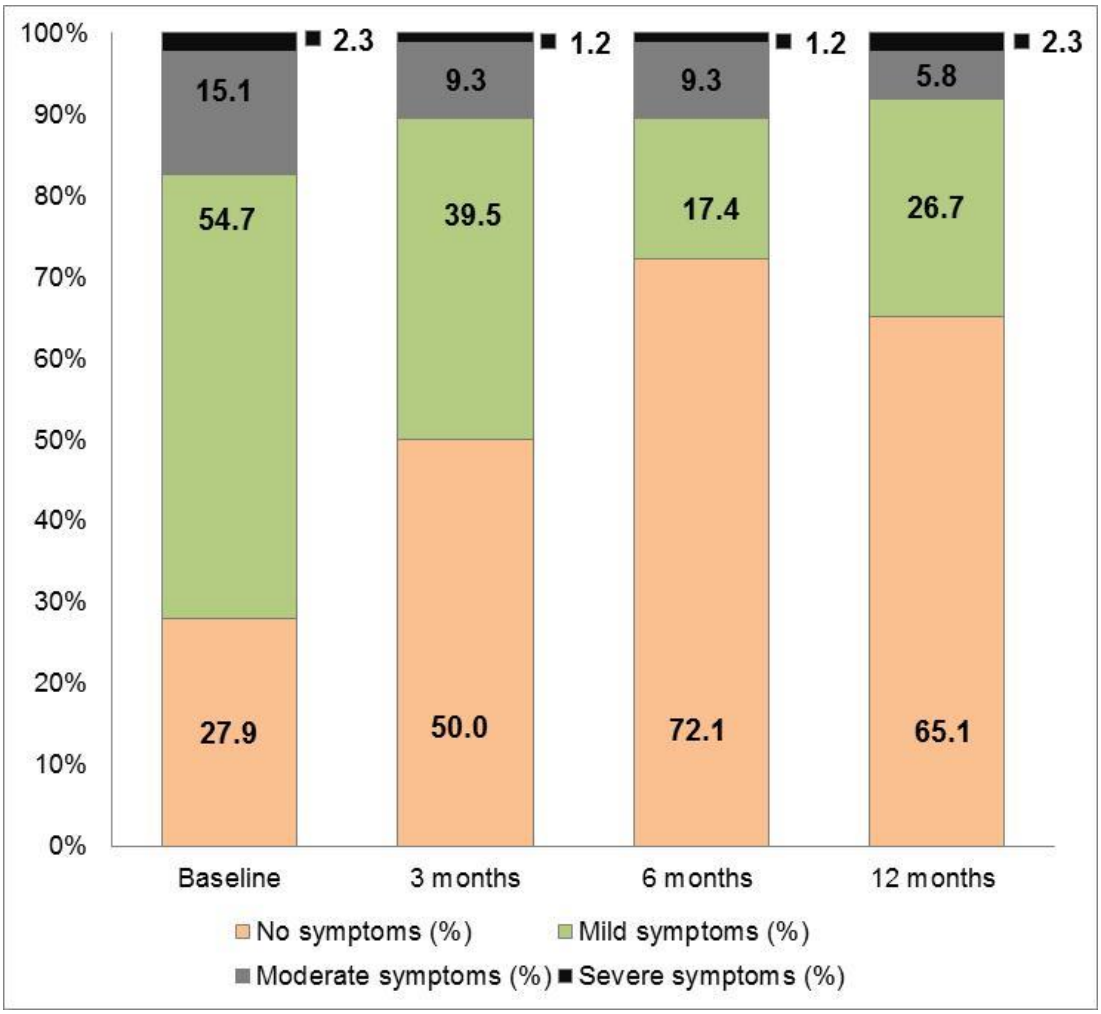
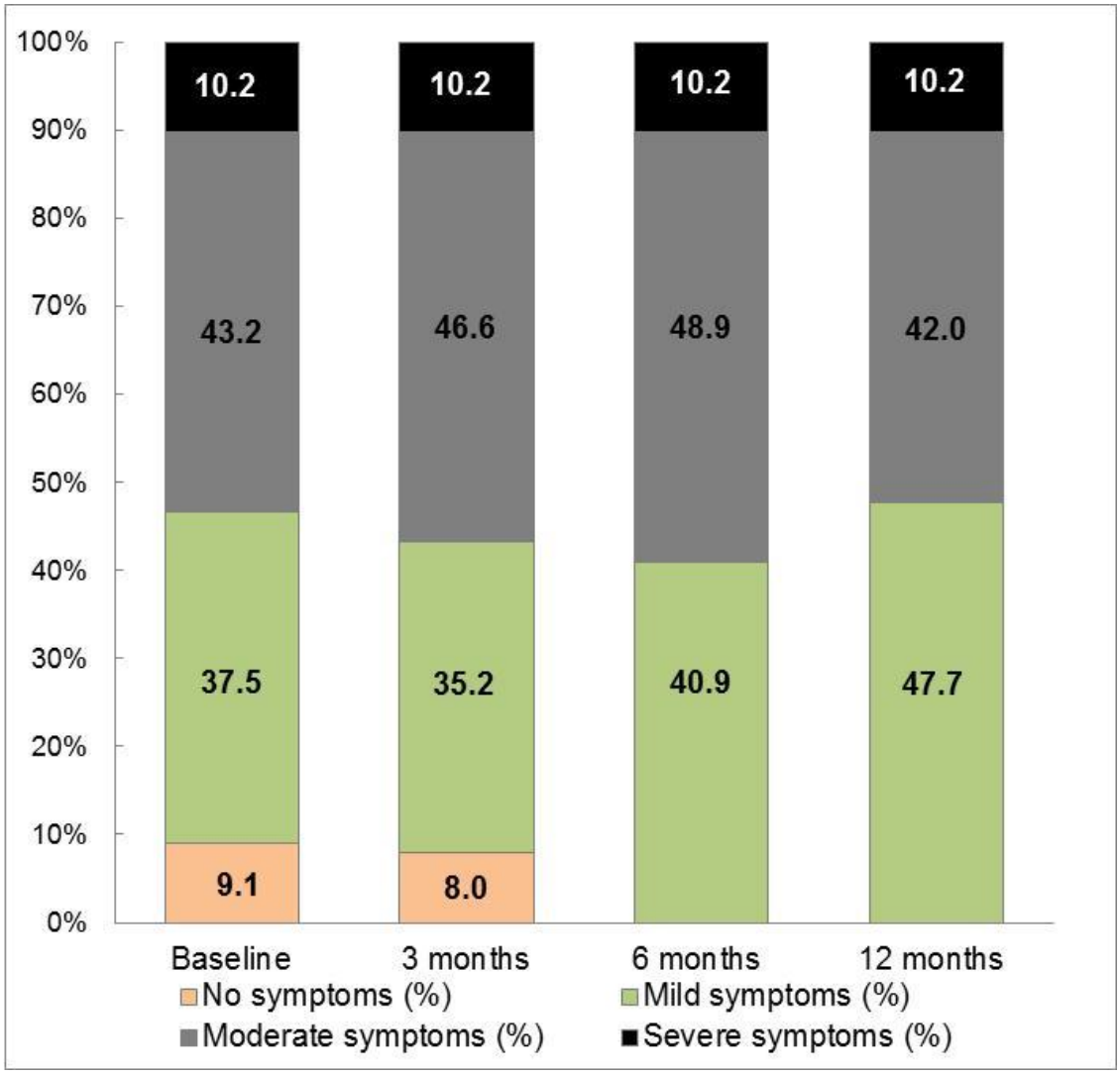


Figure 5.11 on the next page shows the frequencies of persons with the *persistent anxiety symptom trajectory* with different classifications of symptom severity across four time points. 'Mild or worse' anxiety symptoms were reported in all but 8 persons at baseline and 7 persons at 3 months. Approximately half of patients with anxiety had moderate or severe symptom at baseline (n=47), 3 months (n=50), 6 months (n=52) and 12 months (n=46). Figure D.5.3 (in Appendix D.5 on page 383) presents individual HADS-anxiety scores over time for participants in this cluster.

Figure 5.11 3-cluster LCGA anxiety model: *persistent anxiety symptom trajectory* (n=88).



5.7.7 Observed patterns of associations between anxiety and depression symptom trajectories

In total, 91.6% of those with the *no anxiety symptom trajectory* and 80.2% of those with the *transient anxiety symptom trajectory* respectively, had the *no depression symptom trajectory*. Of those in the *persistent anxiety symptom trajectory* 46.6% also were classified to the *persistent depression symptom trajectory*. Together, participants with persistent anxiety symptoms were more

likely to have persistent depression symptoms than individuals with either of two other anxiety trajectories. Table 5.12 shows the frequencies of the *no depression symptom trajectory and the persistent depression symptom trajectory* across persons with the three anxiety trajectories.

Table 5.12 Frequencies of the two depression trajectories across the three trajectories nested in the 3-cluster LCGA anxiety model.

	3-cluster LCGA anxiety model		
2-cluster LCGA depression model	No anxiety symptoms n=119 n (%)	Transient anxiety symptoms n=86 n (%)	Persistent anxiety symptoms n=88 n (%)
No depression symptoms n=232	10 (91.6)	69 (80.2)	46 (52.3)
Persistent depression symptoms n=66	8 (6.7)	14 (16.3)	41 (46.6)

Note: Due to missing data: 6 participants from the LCGA anxiety model had no depression trajectory assigned, 11 participants from the LCGA depression model had no anxiety trajectory assigned.

5.8 DISCUSSION

5.8.1 Summary of key findings

Depression course in older primary care patients consulting with musculoskeletal pain

Overall, 10% of people without depression symptoms (i.e. HADS score 0-7) at baseline reported having elevated depression symptoms at the 1-year follow-up. In total, more than half of participants with elevated depression symptoms at baseline still had depression symptoms at the 1-year follow-up. An analysis of discrete person-centred trajectories of depression symptoms showed that 63% of

older patients with musculoskeletal pain, who were depressed at baseline, had persistent depression symptoms for at least one year. In total, two distinct longitudinal trajectories were identified; 77.8% of patients were classified as having the *no depression symptom trajectory* and 22.2% had the *persistent depression symptom trajectory*. Across four time points, between 35% and 42% of patients with the *persistent depression symptom trajectory*, reported 'moderate or worse' depression symptoms (HADS-D score ≥11).

Anxiety course in older primary care patients consulting with musculoskeletal pain

Only 11% of people without anxiety symptoms (HADS-A score 0-7) at baseline reported elevated anxiety symptoms at the 1-year follow-up. In total, over half of participants with elevated anxiety symptoms at baseline still had symptoms at the 1-year follow-up. In contrast with depression symptom trajectories, finding the optimal model of anxiety symptom trajectories posed some challenges. The LCGA anxiety model with three distinct trajectories was selected as being the best model. This model, included: *no anxiety symptom* (41%), *transient anxiety symptoms* (29%) *persistent anxiety symptom* (30%) *trajectories*. Overall, 56% of older patients with musculoskeletal pain, who were anxious at baseline, had persistent anxiety symptoms for at least one year. At each time point half of the participants with in the *persistent anxiety symptom trajectory* reported 'moderate or worse' anxiety symptoms. Participants with transient anxiety symptoms had typically mild anxiety symptoms at baseline, followed by a symptom decrease or fluctuation at follow-ups. Persons with persistent anxiety symptoms were more likely to have coexisting persistent depression symptoms.

5.8.2 Comparison with previous research

Course of depressive and anxiety symptoms

Comparison with previous research is challenging as the majority of reported estimates of symptom persistence are based on two repeated measures and use variable-centred methods of analyses. In anticipation of this problem, latent class growth analyses were complemented with analyses of changes in the rate of HADS-anxiety and HADS-depression symptoms over time.

Fifty nine percent of older primary care attendees with musculoskeletal pain and coexisting 'mild or worse' depression symptoms at baseline still had depression at the 1-year follow-up. This proportion is broadly comparable to the estimates previously reported for older adults in the community. For older adults depressed at baseline, prevalence rates of those continuing depressed at 1 year was 63.2% (Prince et al.,1998), at 2 years was 61.2% (Harris et al., 2006) and at 3 years were 51.7% (Schoevers et al., 2003) and 50.4% (Beekman et al., 2001). In one study, 35% of older primary care patients recovered from major depression at 1 year (Licht-Strunk et al., 2009a). This prevalence rate was comparable to recovery rate from 'moderate or worse' depression symptoms (34.4%) found in the current study.

Likewise, observed changes in the rate of anxiety symptoms over time are broadly comparable to results of previous relevant studies in older adults in the community. In the current study, 71% of patients with elevated anxiety symptoms at baseline remained anxious at 1 year, and 11% of participants who were not anxious at baseline had elevated anxiety symptoms at 1 year. These estimates are comparable with those reported for older community-dwelling adults (De Beurs et al., 2000). More specifically, 69% of participants found to have persistent anxiety

symptoms at a 3-year and 14.5% of participants who were not anxious at baseline found to have elevated anxiety symptoms at 3 years (De Beurs et al., 2000).

Due to a lack of research, comparisons for anxiety trajectories were not possible. However, the results of latent class growth analyses of depression symptoms can be compared with Licht-Strunk's (2008) study in older general practice attendees. They used a comparable methodology to LCGA -longitudinal latent class analyses (LLCA). In contrast to the current study, it consisted exclusively of patients with depressive disorders at baseline and LLCA analyses were based on raw depression symptoms scores. A recovering group (42%), without elevated symptoms after 6 months follow-up was identified. In addition, three groups of patients with continuous elevated depression symptoms, of different severity mean scores (mild (35%), moderate (18%), severe (5%)). Putting on one side differences between the two studies, arguably they both identified two fairly stable groups of trajectories (characterised by elevated symptoms presence or their lack). Based on baseline depressive score, 63% of older primary care attendees with musculoskeletal pain had elevated depression symptoms over time. A similar frequency (60%) of persistent elevated depression symptoms over time was found in older primary care attendees by Licht-Strunk's (2008). Licht-Strunk (2008) found through univariable multinomial logistic regression analyses that the presence of a chronic somatic illness increases the risk of moderate or severe chronic symptoms, but no differences between the presence of one or more somatic co-morbidity emerged. The impact of the existence of one somatic comorbidity was found to be no longer significant when entered into a multivariable multinomial logistic regression. However, this could be partially affected by inclusion of as many as 11 factors in comparison of relatively small groups (125,

104 and 67 participants). Consequently, the impact of OA presence on depressive and anxiety symptoms persistence in older adults still warrants clarification.

5.8.3 Strengths and limitations

Strengths

This investigation has two major strengths that go beyond the strength of the data source used here (i.e. PROG-RES study that was delineated in chapter four). A key strength is the relevance to primary care. This study offered the specific attention to anxiety symptoms, which to date has been under-research and overlooked in primary care guidance. Analyses focused on symptom severity, as in UK primary care, depression symptom severity plays an important part in the choice of intervention (NICE, 2009b). Furthermore, anxiety and depression symptoms were ascertained with the HADS, which is a standardised questionnaire recommended for primary care usage (discussed in chapter three).

The research method is also a major strength of this study. The analyses of depressive and anxiety symptoms over time went beyond the usual description of the course of prevalence rates by presenting an in-depth exploration of trajectories. The identification of trajectories involved the use of a standardised person-centred method that is tailored to longitudinal data analyses and relatively easy to interpret.

Limitations

As highlighted by Shuurmans et al. (2005) in the context of the course of anxiety disorder in older adults, dichotomisation of the main outcome can lead to the loss of information. Nevertheless, the dichotomisation can be argued to make

data summarisation more efficient, and allows for simple interpretation of results. In the current study dichotomisation of HADS depression and anxiety scores simplified the task of grouping and describing observed trajectories, in each discrete trajectory. It can be argued that dichotomisation is appropriate only when a threshold effect value is meaningful (Abdollel et al., 2002). To date, biomarkers for clinically significant depressive and anxiety symptoms are lacking and diagnostic accuracy of all depressive and anxiety questionnaires is widely criticised. Consequently, it seems that meaningfulness of cut-off points on any depression and anxiety questionnaires is unlikely to represent 'truly' existing thresholds. However, given that NICE depression guidance for patients with a chronic physical health problem refers to mild depression symptoms (NICE, 2009b), using HADS score ≥8 was a justified choice.

5.8.4 Research Implications

There are several research implications arising from this study. Validation analyses would be beneficial, particularly for the anxiety model, using a larger sample size and software such as MPlus 3, which is less user-friendly than the Latent GOLD software, but promises some useful improvements in the estimation of LCGA models (Nagin & Tremblay, 2001).

As demonstrated in chapter four, the sample included in this analysis is comprised of patients with a range of symptoms, with those with mild levels of disability and those with new episodes of depressive and anxiety symptoms underrepresented. As a result, the prevalence rates of individuals with emerging symptom trajectories could be underestimated. Future research would benefit from

performing analyses stratified by the length of endured pain with balanced proportions of new episodes and recurring musculoskeletal pain.

GMM, LLCA and LCGA allow for the prediction of cluster membership by adding covariates to the model. Whilst this can add information, entering covariates to the model can substantially change the parameters of the trajectories of the outcome variable (Feldman et al, 2009). Nylund and Masyn (2008) (cited in Feldman et al., 2009) suggest that the number of clusters should be determined using an unconditional model (without covariates). Feldman et al. (2009) suggest that entering covariates should be based on a strong theoretical basis, with careful consideration of the interpretability and parsimony of the model. As such, a limited number of covariates should be entered, yet no specific number has been suggested. Consequently, including covariates was deemed unsuitable for the aim of this study, which intended to explore patterns of anxiety and depression. The next logical step in analyses of trajectories identification of person-related characteristics associated with different trajectories (e.g. Croudce et al., 2003, Licht-Strunk, 2008, Olino et al., 2010). This type of study has a potential to inform mechanisms underlying coexistence of OA and problems of depression and anxiety. Identified characteristics may also inform future research in targeting patients at risk of poor anxiety depression prognosis.

5.9 CONCLUSIONS

The current study is the first exploration of the course of depressive and anxiety symptoms in older patients consulting primary with musculoskeletal pain. Two discrete trajectories of depression symptoms over the 12-month period were identified, namely *no depression symptom* (78%) and *persistent depression*

symptom (22%) trajectories. Three anxiety trajectories were identified, including *no anxiety symptom* (41%), *transient anxiety symptom* (29%) *and persistent anxiety symptom* (30%) *trajectories*. With the aim of identifying cases with persistent depressive or anxiety problems in people who have mild to severe depressive and anxiety symptoms at the initial presentation, it can be anticipated that 63% and 56% of these individual will have persistent depressive and anxiety symptoms respectively over the 12-month period.

Older people with musculoskeletal pain experiencing persistent symptoms of anxiety and depression are most likely to be suffering adverse effects of depressive and anxiety symptomology. This highlights the importance of enquiring about addressing these problems in persons with OA. Further research is needed to understand factors associated with trajectories of anxiety and depression symptoms.

Chapter six: The course of anxiety and depression symptoms in older patients presenting to general practice with musculoskeletal pain

Part 3: Factors associated with the course of anxiety and depression symptoms

6.1 INTRODUCTION

Chapter five described the course of depression and anxiety symptoms. This was determined based on the HADS questionnaire over the 12-month period in older patients presenting to general practice with musculoskeletal pain. Three anxiety and two depression symptom trajectories were identified. As described in research applications of the previous chapter (p. 199), this chapter includes the next logical step following identification of these trajectories. The identified growth trajectories are analysed for associations with person-related characteristics, using the baseline covariate data described in section 4.4.5 (p. 147).

6.2 RATIONALE OF THE STUDY

In addition to the fear-anxiety avoidance (Asmundson et al., 2004), misdirected-problem solving (Eccleston & Crombez, 2007) and the acceptance-commitment (Hayes et al., 1999) models, this thesis contends that coexisting anxiety and depression might reflect the multifactorial process of adjustment to OA symptoms. Empirical application of these theoretical concepts in subpopulations of people with musculoskeletal pain or OA specifically highlight the importance of pain severity, pain impact on daily functioning and coping strategies (Hayes et al., 1999, De Vlieger et al., 2006, Kratz et al., 2007, Scopaz et al., 2009, Flink et al., 2011). Lee and Mercurio-Riley's (2009) review

of the contributing factors in psychosocial adjustment to chronic pain of musculoskeletal origin, identified as many as six groups of associated factors. Factors included *pain conditions*, *functional dependence*, *stress*, *stress processing*, *intrapersonal factors* and *socio-ecological factors* (see Figure 1.1 on page 15 for more specific examples). Less is known about which of these factors can predict poor depressive and anxiety prognosis in depressed/anxious adults with musculoskeletal pain. Cole et al.'s (1999) review suggests that a number of factors have been analysed for their association with poor prognosis of depression at 24 months in elderly community, where no factors appeared to be predominately associated with poor depression outcome. Licht-Strunk et al. (2007) concluded in a comprehensive systematic review (including assessment of the strength of association) found that general practice evidence provides no strong support for any predictors of a poor depression outcome. "*In community studies strong evidence [i.e. defined as significant associations with poor outcome in at least two high-quality cohorts] was found for older age, the presence of chronic somatic diseases, the presence of functional limitations, higher baseline depression level and an external locus of control*" (Licht-Strunk et al., p. 172, 2007). To the best knowledge of the author of this thesis, there is no review of predictors of the curse of anxiety in elderly community or primary care.

The previous chapter described the first exploration of the course of depressive and anxiety symptoms in older primary care patients with musculoskeletal pain. Inclusion of patient-related characteristics in developed models was deemed methodologically unsuitable for the aim of that study. The next logical step, however, is a need to identify person-related characteristics associated with different trajectories, particularly with persistent depressive and

anxiety symptoms. This type of study has two important implications. Firstly, it has the potential to inform mechanisms underlying the coexistence of OA and depression and anxiety responses. This might be used to inform the usefulness of psychological interventions that target mechanisms underlying adjustment to pain. The second important implication is that identified characteristics may also inform future research in targeting patients at risk of poor anxiety depression prognosis. For example, according to Licht-Strunk et al. (2009a), identifying patients at a high risk of persistent depressive and anxiety symptoms may improve their management and subsequent prognosis. This is also supported by clinical guidelines for depression that recommend for the recognition of patients with a chronic physical health problem a need of depression management to be based on contextual factors (NICE, 2009b). To date limited information is available to support recognition of the problem of depression in the community-dwelling patients with OA and improving the situation for anxiety seems even more demanding. The analysis presented in this chapter will investigate person-related characteristics associated with the course of anxiety and depression symptoms in older people with musculoskeletal pain in primary care.

6.3 AIM AND OBJECTIVE

The **overall aim** of this study is to advance understanding of the persistence of depressive and anxiety symptoms in older primary care patients with OA.

Specific objective is:

- To examine their unique relationships of the course of anxiety and depression symptoms with baseline person-related characteristics

6.4 METHOD

6.4.1 Sample

Analyses of factors associated with the identified trajectories were based on samples with at least three HADS depression scores and anxiety scores (i.e. at 3 time points). This decision was grounded in statistical reasoning, that is, to improve the statistical power of regression analyses by increasing sample sizes. Selected were only participants with at least three HADS scores to avoid potential misclassifications of cluster membership, as minimum three repeated measures are required for conducting LCGA (Nagin, 2005, Vermunt & Magidson, 2005, Peng, 2011).

6.4.2 Statistical analyses

Each participant was assigned a cluster membership by re-estimating the selected optimal models (i.e. the 2-cluster LCGA depression model and the 3-cluster LCGA anxiety model). The Latent GOLD handled missing data in the likelihood fashion, i.e. models were estimated using full-information maximum likelihood, which uses all available data and assumes that data are missing at random (Vermunt & Magdison, 2005). Characteristics of the trajectories of depressive and anxiety symptoms were then compared with the reference models.

To explore factors associated with trajectories, in line with previous studies (e.g. Olino et al., 2010), descriptive statistics and logistic regression was used to compare the baseline characteristics of participants with different anxiety and depression trajectories. The baseline characteristics of interest were discussed in details in chapter four (section 4.4.5 on page 147). A brief summary is provided in Table 6.1 overleaf.

Table 6.1 Baseline variables of interest.

Variables	Measure	Time frame	Type	Format used in logistic regression analyses
Age	Date of birth	≥ 50 years	ordinal	1=50-59 years 2=60-69 years 3=70+ years
Availability of support: Emotional Instrumental	1 item (yes/no/no need)	not specified	categorical	0=yes/no need 1=no
Coping strategies: Catastrophising Increased behavioural activities Coping self-statements Ignoring pains sensations	II-items version of the CSQ (the average of 2 items)	never-always	categorical	below vs. above the highest tertile 0=score <4 1=score ≥4 0=score < 5 1=score ≥5 0=score <4.5 1=score ≥4.5 0=score <4 1=score ≥4
Gender	Female/male	N/A	categorical	0=male 1=female
Living arrangement	Living alone (yes/no)	currently	categorical	0=not living alone 1=living alone
NS-SEC classification	Job title	current or most resent	categorical	0=managerial and professional/ intermediate/other 1=routine and manual
Number of pain sites	Manikin	past 4 weeks	count	0-44 pain sites
Pain interference with: a) daily activities b) work/housework c) social and family activities	1 item each on Chronic pain Grade	past 3 months	continuous	0-10 NRS
Marital status	Married/ single/ divorced/ widowed/ separated/ cohabiting	currently	categorical	0=married/cohabiting 1=divorced/widowed/ separated/single

Note: CSQ- Coping Strategies Questionnaire; NRS- numerical rating scale; NS-SEC- National Statistics-Socioeconomic Class 2010 classification.

Where the LCGA analysis suggested a 2-cluster solution was the best fit (i.e. the 2-cluster depression model), binary logistic regression ($r \leq 2$ categories) was used to compare participants in each of the two trajectories. Where the LCGA solution included more than 2 clusters (i.e. the 3-cluster anxiety model), multinomial logistic regression was used, which can handle polytomous responses ($r > 2$ categories). STATA version 11.1 was used for both logistic regression (command *logit*) and multinomial logistic regression (command *mlogit, rrr*).

Regression analyses commenced with the preliminary identification of a potential problem of multicollinearity between analysed covariates (i.e. a highly correlated predictors that provide redundant information about the response). This involved conducting a variance inflation factor (VIF) analysis in STATA (command *vif*). For each factor tolerance and VIF estimates were estimated, where a tolerance of less than 0.10 and a VIF more than 10 (or ≤0.20 and ≥5.0 respectively) indicated a possibility of multicollinearity (O'Brien et al., 2007). A univariate regression analysis was then conducted and variables with apparently wide 95% CI were identified and excluded. Following this, univariate regression analyses were conducted with and without a variable suspected of causing multicollinearity and changes in odds ratios and 95% CI of the remaining covariates were examined. If multicollinearity was confirmed the variable was excluded from further regression analyses.

In order to preserve a model's parsimony (Nagin, 2005) for the remaining covariates a method a backward elimination regression analyses were used with 0.10 entry probability based on the Wald test. Statistical significance (for variables that met the entry probability) was established at the level of $p<0.05$. Backward elimination logistic regression analyses were used for the 2-cluster LCGA

depression model, with the *no depression symptom trajectory* as a reference group. The 3-cluster LCGA anxiety model was analysed using backward elimination multinomial logistic regression. Analyses commenced with the *no anxiety symptom trajectory* acting as a reference group, then the model was re-run using the *persistent anxiety symptom trajectory* as a reference category. To maintain consistency, odds ratios for both logistic and multinomial logistic regression were reported. As STATA allows for calculating relative risk ratios (RRRs) but not ORs, STATA technical guideline (Gould, 2000) was searched for solutions and e-mail communication with STATA support team was made. Both suggested that RRRs can be used as ORs. To confirm it multinomial logistic regressions in PASW version 18.0 were conducted, which provides estimates of ORs for multinomial logistic regression. Similar estimates were produced by sets of statistical software.

6.5 RESULTS

6.5.1 Sample size

In total, 368 participants (83% of consenters for follow-up) had at least three HADS depression and anxiety scores recorded. Average posterior probabilities and cluster characteristics of the 2-cluster LCGA depression and the 3-cluster LCGA anxiety models are reported in Appendices E.3 (on page 387) and E.4 (on page 388) respectively.

6.5.2 The issue of multicollinearity in regression analyses

The result of the VIF test for multicollinearity suggests possible co-linearity with one of the factors. The ‘pain interference with daily activities’ variable resulted

in a VIF of 5.19 (see Table E.1.1 in Appendix E.1 on page 384), suggesting a need for further exploration of the impact of this variable on the models. A problem of multicollinearity was detected for the pain interference with daily activities variable, as indicated by changes in ORs after excluding this variable from saturated logistic regression and multinomial logistic regression analyses. Furthermore, the lack of emotional support variable had wide confidence intervals (95% CI 0.51, 29.7 and 1.27, 27.38) (see Tables E.2.1 and E.2.2 in Appendix E.2 on pages 385-386), related to a small number of people who reported a lack of emotional support (n=16). Therefore, although both variables were included in descriptive analyses, they were excluded from backward elimination logistic regression and backward elimination multinomial logistic regression analyses.

6.5.3 Baseline factors associated with the depression symptom trajectories

Descriptive statistics

Demographic and clinical characteristics at baseline for the 368 patients with at least three HADS-D scores, split by the *no depression symptom trajectory* (n=272) and the *persistent depression symptom trajectory* (n=96) are shown in Table 6.2 overleaf. Participants with persistent depression symptoms appear to have more widespread pain and pain interference with work, social activities and daily activities, when compared to participants in the *no depression symptom trajectory*. The *persistent depression symptom trajectory* included higher proportions of participants: with perceived lack of emotional and instrumental supports, coping by catastrophising, females, the oldest adults, people without a partner, living alone, performing routine or manual work. Participants in the *no*

depression symptom trajectory had higher frequency of people coping by self-statements, ignoring sensations and increased behavioural activities.

Table 6.2 Individual characteristics across people with no depression symptoms and persistent depression symptoms (n= 368).

Baseline covariates	No depression symptoms n= 272, n (%)	Persistent depression symptoms n= 96, n (%)
Age: 50-59	102 (37.5)	25 (26.0)
60-69	101 (37.1)	26 (27.1)
70+	69 (25.4)	45 (46.9)
Gender:		
Females	163 (59.9)	63 (65.6)
Males	109 (40.1)	33 (34.4)
Lack of partner		
Married/cohabiting	215 (79.0)	68 (70.8)
Single/divorced/widowed/separated	55 (20.2)	27 (28.1)
Number of pain sites Median (IQR)	6.0 (8)	10.5 (14)
Pain interference with activities: Median (IQR)		
Social and family activities	5.0 (5)	8.0 (3)
Daily activities	5.0 (4)	8.0 (2)
Work	4.0 (4)	8.0 (3)
Living alone		
Yes	41 (15.1)	21 (21.9)
No	230 (84.6)	75 (78.1)
Catastrophising†		
Not high	196 (72.1)	42 (43.8)
High	60 (22.1)	45 (46.9)
Coping by increased behavioural activities†		
Not high	114 (41.9)	45 (46.9)
High	147 (54.0)	38 (39.6)
Coping by ignoring pain sensations†		
Not high	187(68.8)	57 (59.4)
High	71 (26.1)	24 (25.0)
Coping by using self-statement†		
Not high	137 (50.4)	53 (55.2)
High	126 (46.3)	33 (34.4)
Availability of emotional support		
Yes/no need	265 (97.4)	84 (87.5)
No	4.0 (1.5)	12.0 (12.5)
Availability of instrumental support		
Yes/no need	255 (93.8)	78 (81.3)
No	16 (5.9)	18 (18.8)
Manual/ routine work		
No	175 (64.3)	59 (61.5)
Yes	97 (35.7)	37 (38.5)

Note: IQR- inter quartile range;
Percentage does not always add up to 100% due to missing data;
†- Subscales from 2-item Coping strategy Questionnaires (Jensen et al., 2003). Cut-offs represent heights tertile in the current sample (see p. 206).

Logistic regression analyses

Fourteen covariates were entered into the backward elimination model with a total number of 314 participants included in the analysis. Among the variables that failed to meet p<0.10 entry probability were: living alone (p=0.819), gender (p=0.742), coping by using self-statements (p=0.753) and ignoring pain sensations (p=0.762), the manual/routine work class (p=0.479), pain interference with work (p=0.461), a lack of partner (p=0.379), age 60-69 (p=0.291).

Variables that met the entry probability threshold are listed in Table 6.3 overleaf. Relative to the *no depression symptoms trajectory*, for every unit increase in pain interference with social activities the risk of having the *persistent depression trajectory* increased by 30.0%. With pain in one additional anatomical site, the risk of having persistent depression symptoms increased by 8.0%. Patients age 70 years or above demonstrated 3 times the risk of having the *persistent depression symptom trajectory*, as opposed to the *no depression symptom trajectory*. Pain coping with increased behavioural activities was protective of persistent depression symptoms (OR: 0.51). Participants with a lack of instrumental support had 4 times the risk of being in the *persistent depression symptom trajectory*, but wide confidence intervals indicate that the impact should be interpreted with caution.

Table 6.3 Backward elimination logistic regression analysis of baseline covariates: *no depression symptom*^ vs. *persistent depression symptom trajectories*.

Baseline covariates	Odds Ratio	SE	p	95% CI	
Age 70 years or older	3.03	1.01	0.001	1.57	5.84
Interference with social activities (0-10)	1.30	0.09	<0.0001	1.14	1.48
Lack of instrumental support	3.63	1.87	0.013	1.32	9.98
Number of pain sites (0-44)	1.08	0.02	<0.0001	1.04	1.13
Catastrophising	1.85	0.62	0.066	0.96	3.55
Coping by increased behavioural activities	0.51	0.16	0.037	0.27	0.96

Note: CI- confidence intervals; SE- standard error; ^- the reference group;
The model was statistically significant (LL= -127.77, LR chi^2 (6) = 85.00, Prob > chi p<0.001, Pseudo R^2= 0.25).

6.5.4 Baseline factors associated with the anxiety symptom trajectories

Descriptive statistics

Table 6.4 Individual characteristics across three anxiety trajectories (n= 368).

Baseline covariates	No anxiety symptoms n=142 n (%)	Persistent anxiety symptoms n=123 n (%)	Transient anxiety symptoms n=103 n (%)
Age: 50-59	45 (31.7)	53 (43.1)	31 (30.1)
60-69	58 (40.8)	30 (24.4)	38 (36.9)
70+	39 (27.5)	40 (32.5)	34 (33.0)
Gender:			
Males	64 (45.1)	39 (31.7)	38 (36.9)
Females	78 (54.9)	84 (68.3)	65 (63.1)
Lack of partner			
Married/cohabiting	113 (79.6)	92 (74.8)	79 (76.7)
Single/divorced/widowed/ separated	28 (19.7)	30 (24.4)	23 (22.3)
Number of pain sites Median (IQR)	5.5 (6.0)	9.0 (12.0)	8.0 (10.0)
Pain interference with activities: Median (IQR)			
Social and family activities	4.0 (5.0)	7.0 (4.0)	6.0 (5.0)
Daily activities	4.0 (5.0)	7.0 (3.0)	6.0 (4.0)
Work	3.0 (5.0)	6.0 (4.0)	6.0 (4.0)
Living alone			
Yes	21 (14.8)	23 (18.7)	16 (15.5)
No	120 (84.5)	100 (81.3)	87 (84.5)
Catastrophising†			
Not high	119 (83.8)	55 (44.7)	63 (61.2)
High	14 (9.9)	56 (45.5)	36 (35.0)
Coping by increased behavioural activities†			
Not high	53 (37.3)	60 (48.8)	49 (47.6)
High	80 (56.3)	50 (40.7)	51 (49.5)
Coping by ignoring pain sensations†			
Not high	99 (69.7)	76 (61.8)	69 (67.0)
High	34 (23.9)	31 (25.2)	28 (27.2)
Coping by using self-statement†			
Not high	70 (49.3)	67 (54.4)	54 (52.4)
High	66 (46.5)	45 (36.6)	45 (43.7)
Availability of emotional support			
Yes/no need	139 (97.9)	110 (89.4)	100 (97.1)
No	1 (0.7)	12 (9.8)	3 (2.9)
Availability of instrumental support			
Yes/no need	137 (96.5)	102 (82.9)	95 (92.2)
No	4 (2.8)	21 (17.1)	8 (7.8)
Manual/ routine work:			
No	100 (70.4)	78 (63.4)	54 (52.4)
Yes	42 (29.6)	45 (36.6)	49 (47.6)

Note: IQR- inter quartile range;
Percentage does not always add up to 100% due to missing data;
†- Subscales from 2-item Coping strategy Questionnaires (Jensen et al., 2003). Cut-offs represent heights tertile in the current sample (see p. 206).

Table 6.4 on the previous page shows characteristics at baseline for the 368 patients with at least three HADS-A scores, stratified by people with no anxiety symptoms (n=142), persistent anxiety symptoms (n=123), transient anxiety symptoms (n=103).

By comparison with other trajectories, patients with the *no anxiety symptom trajectory* had: a lower number of pain sites, lower levels of pain interferences with activities, lower frequency of coping by catastrophising, higher frequency of coping with increased behavioural activities and a lower proportion of people performing manual/routine work. Individuals within the *no anxiety symptom trajectory* had lower proportions of the youngest patients and those without support than people with persistent anxiety symptoms. The *no anxiety symptom trajectory* was characterised by a higher proportion of males than the *persistent anxiety symptom trajectory*.

Contrasting the *transient anxiety symptom trajectory* with the *persistent anxiety symptom trajectory*, suggest that participants with the latter: more often were in the youngest and the oldest age groups, were more often females, had a higher number of pain sites, higher levels of pain interference with daily and social activities, more often lived alone, more often coped by catastrophising, less frequently coped by increased behavioural activities and self-statements, had less perceived support and less frequently performed manual/routine work.

Multinomial logistic regression analyses

Fourteen covariates were entered into a backward elimination multinomial regression model, where the *no anxiety symptom trajectory* served as a reference group. A total number of 316 persons were included in the analysis. Variables that

did not meet the 0.10 entry probability included: living alone (p=0.997), coping with self-statements (p=0.886), a lack of a partner (p=0.847), pain interference with social activities (p=0.832), age 70 years or above (p=0.541), gender (p=0.387), ignoring pain sensations (p=0.223), and age 60-69 (p=0.113).

Table 6.5 overleaf shows the results of a comparison between persons with no anxiety symptoms and persistent anxiety symptoms. Catastrophising and pain interference with work were the most prominent covariates. Catastrophising increased the risk of the *persistent anxiety symptom trajectory* (OR: 4.14). For every unit increase in the pain interference with work variable the odds of having the *persistent anxiety symptom trajectory* increased by 21%. With pain in one additional anatomical site, the odds of having the *persistent anxiety symptom trajectory* increased by 8%. Pain coping with increased behavioural activities decreased the risk of having persistent anxiety symptoms (as indicated by an odds ratio 0.40). Individuals reporting a lack of instrumental support had seven times the risk of having persistent anxiety symptoms, but a wide confidence interval indicates that more data needs to be collected to confirm the impact of this covariate.

Table 6.5 Backward elimination multinomial logistic regression analysis of baseline covariates: *no anxiety symptom^* vs. *persistent anxiety symptom trajectories*.

Baseline covariates	Odds Ratio	SE	p	95% CI	
Catastrophising	4.14	1.61	<0.0001	1.93	8.87
Coping by increased behavioural activities	0.40	0.13	0.004	0.22	0.75
Lack of instrumental support	6.99	4.93	0.006	1.75	27.85
Manual/routine work	1.33	0.44	0.387	0.70	2.54
Number of pain sites (0-44)	1.08	0.03	0.001	1.03	1.14
Pain interference with work (0-10)	1.21	0.07	0.001	1.08	1.36

Note: CI- confidence intervals; SE- standard error; ^- the reference group;
The model was statistically significant (LL=-342.39, LR chi^2 (18) = 106.42, Prob > chi p<0.001, Pseudo R^2=0.13).

Table 6.6 overleaf shows the results of a comparison between participants with no anxiety symptoms and transient anxiety symptoms. For every unit increase in the pain interference with work variable the odds of having the *transient anxiety symptom trajectory* increased by 19%. Catastrophising increased the likelihood of the *transient anxiety symptom trajectory*, being reflected in an odds ratio 2.79. Performing manual/routine work is a prominent factor, increasing the likelihood of having transient anxiety symptoms (as reflected in an odds ratio 2.39, 95% CI 1.31, 7.60, p=0.005). With pain in one additional anatomical site, the odds of having the *persistent anxiety symptom trajectory* increased by 6%. Pain coping with increased behavioural activities had a borderline effect on cluster membership.

Table 6.6 Backward elimination multinomial logistic regression analysis of baseline covariates: *no anxiety symptom^* vs. *transient anxiety symptom trajectories*.

Baseline covariates	Odds Ratio	SE	p	95% CI	
Catastrophising	2.79	1.08	0.008	1.30	6.00
Coping by increased behavioural activities	0.55	0.17	0.050	0.31	1.00
Lack of instrumental support	1.54	1.25	0.598	0.31	7.60
Manual/routine work	2.39	0.74	0.005	1.31	4.37
Number of pain sites (0-44)	1.06	0.03	0.013	1.01	1.11
Pain interference with work (0-10)	1.19	0.07	0.002	1.06	1.33

Note: The model was statistically significant (LL=-342.39, LR chi^2 (18) = 106.42, Prob > chi p<0.001, Pseudo R^2=0.13); ^- the reference group'; SE- standard error; CI- confidence intervals.

As described in section 6.4.2 (on page 205), backward elimination multinomial regression analyses were repeated using the same variables, with the *persistent anxiety symptom trajectory* serving as a reference category. Table 6.7 presents the results of comparisons with *the transient anxiety symptom trajectory*. No significant differences were found between the two groups. Preforming manual/routine activities had a borderline effect on cluster membership, with odds ratio suggesting a possibility of manual/routine increasing the likelihood of transient anxiety symptoms.

Table 6.7 Backward elimination multinomial logistic regression analysis of baseline covariates: *persistent anxiety symptom^* vs. *transient anxiety symptom trajectories*.

Baseline covariates	Odds Ratio	SE	p	95% CI	
Catastrophising	0.67	0.22	0.281	0.77	2.43
Coping by increased behavioural activities	1.37	0.40	0.698	0.59	2.21
Lack of instrumental support	0.49	0.24	0.144	0.18	1.28
Manual/routine work	1.80	0.54	0.051	1.00	3.24
Number of pain sites (0-44)	0.98	0.02	0.273	0.94	1.02
Pain interference with work (0-10)	0.98	0.06	0.758	0.88	1.10

Note: CI- confidence intervals; SE- standard error; ^- the reference group;
The model was statistically significant (LL=-342.39, LR chi^2 (18) = 106.42, Prob > chi p<0.001, Pseudo R^2=0.13).

6.6 DISCUSSION

6.6.1 Summary of key findings

In comparison with individuals with the *no depression symptom trajectory*, persons with persistent depression symptoms were more likely to be older, report more widespread pain, have more severe pain interference with social activities, less often cope with their pain through increased behavioural activities and more frequently perceive a lack of instrumental support.

When compared to individuals with the *no anxiety symptom trajectory*, participants with persistent anxiety symptoms were more likely to catastrophise, report more widespread pain and more severe pain interference with work, and less often cope with their pain through increased behavioural activities. They also appeared to perceive a lack of instrumental support, but more data is needed to confirm the impact of this variable. In comparison with individuals with the *no anxiety symptom trajectory*, those with transient anxiety symptoms were more likely to catastrophise, be members of manual or routine occupational class, report more widespread pain and more severe pain interference with work. Persons with persistent anxiety symptoms and transient anxiety symptoms did not significantly differ across any of the explored variables.

6.6.2 Comparison with previous research

No direct causal inferences can be made with regards to the models of adjustment to pain - three of which were introduced in section 1.3.2 (p. 12). Nevertheless, this study is potentially informative for understanding the elements involved in the process of adjustment to musculoskeletal pain in older people.

In the fear-anxiety-avoidance model (Asmundson et al., 2004), pain characteristics

play an important role in the adjustment to pain. This is especially true for initial pain severity and disability, where the latter is preceded by anxiety responses and followed by depressive responses. Indeed, having a higher number of pain sites and higher pain interference with activities are associated with a poorer adjustment to musculoskeletal pain in older adults which is reflected in persistent anxiety and depression symptoms. The current study expands the understanding of the role of specific aspects of disability and depressive and anxiety responses. It suggests that the level of interference with social activities is associated with persistent depression symptoms and levels of interferences with work with persistent anxiety symptoms. The importance of pain characteristics on depressive and anxiety symptoms over time has also been highlighted in the studies by Gerrits et al. (2012) and Norton et al. (2011).

Fear-anxiety-avoidance (Asmundson et al., 2004), misdirected-problem solving (Eccleston & Crombez, 2007) and acceptance-commitment (Hayes et al., 1999) models suggest that coping strategies play a key role in the process of adjustment to pain. In the fear-anxiety-avoidance model having catastrophic thoughts reflects the appraisal of a painful stimulus as a threat, which is likely to result in anxiety and then depressive responses (Asmundson et al., 2004). The current study found that catastrophising is indeed associated with persistent anxiety symptoms, though less strongly associated with persistent depression symptoms. Interestingly, a protective role of increased behavioural activities offers an important insight into difficulties with adjustment to musculoskeletal pain in older adults with musculoskeletal pain, which is reflected by persistent depressive and anxiety symptoms. Misdirected-problems solving (Eccleston & Crombez, 2007) and acceptance-commitment (Hayes et al., 1999) models suggest that people with pain who direct their

efforts to continuing to engage in meaningful life activities (as opposed to focusing on a reduction of pain) are likely to better adjust to pain. It can be argued that older people with musculoskeletal pain who adjusted to pain by focusing on achievable goals (i.e. increased activities), showed better adjustment to pain which is reflected in a decreased chance of persistent anxiety and depression symptoms.

In modern psychological models of pain adjustment (Hayes et al., 1999, Asmundson et al., 2004, Eccleston & Crombez, 2007), socio-ecological aspects seem to be viewed as contextual factors, but their exact roles are unspecified. The impact of socio-ecological factors was limited in older people with musculoskeletal pain, as instrumental support was only found to be weakly associated with persistent anxiety depressive and anxiety symptoms. Emotional support, marital status, living alone and socio-economic status were found to have no significant impact on persistent depression and anxiety trajectories. Interestingly, performing manual/routine work was associated with transient anxiety symptoms. Arguably the risk of symptoms persistence can be expected to be greater for those with lowest employment status. However, evidence suggests that the risk of anxiety is associated with the level of control at work that was found to be unevenly distributed across types of occupations (Griffin et al., 2002). Unfortunately, no information on levels of control in the sample investigated in this thesis is available. However, older people with musculoskeletal pain and transient anxiety symptoms did not significantly differ in coping by increased behavioural activities form individuals without anxiety symptoms over time (albeit the difference was borderline). Previous research found a positive, moderate and significant association between levels of coping by increased behavioural activities and perceived pain control (Haythonthwaite et al., 1998). It could be then argued that a sense of control over pain may

differentiate between transience and persistence of anxiety symptoms, but this hypothesis warrant further investigation.

Older age may serve as a predisposing factor to poor adjustment to pain and there are several plausible explanations for its association with persistent depression symptoms. One possible explanation is that a risk of multimorbidity increases with age (Marengoni et al., 2011). Another possibility is that older people appear to be less likely to consult their GPs for depressive and anxiety problems and therefore have limited access to relevant treatment (RCGP, 2006). The exact reason for the observed association between older age and the persistence of depression symptoms in primary care patients with musculoskeletal pain remains unclear.

6.6.3 Strengths and limitations

Strengths

One strength of this study is a careful selection of variables. Considered were person-related characteristics, which are of potential use to identification of patients with OA and the problem of anxiety or depression. Most of the variables were selected on the basis of being listed in the conceptual framework of risks or resistance factors in adjustment to chronic pain. All of the variables have been considered in modern psychological theorises for their role in adjustment to pain (described in section 1.3.2 on page 12). When feasible, included variables were categorised, to provide clear interpretation of their impacts. The issue of multicoollinearity was addressed. Overall, both implications for primary care practice and statistical parsimony of models and were considered.

In estimating the impact of these variables, this study went beyond purely descriptive approach by using regression analyses. This type of analyses allows

for probabilistic interpretation of the relationship between the likelihood of an event occurring and a set of conditions (Rindskopf, 2004).

Limitations

This study does not allow for making inferences about the impact of intra-personal factors and clinical factors on the trajectories of depressive and anxiety symptoms. Furthermore, only limited conclusions can also be drawn about socio-ecological factors. Examples of intrapersonal factors previously found to be associated with the course of depressive and anxiety disorder symptoms respectively include locus of control (Beekman et al., 2001, Harris et al., 2006) and neuroticism (De Beurs et al., 2000, Schuurmans et al., 2005). Concerning clinical factors, the PROG-RES study was not designed to make inferences about treatment effectiveness. Subsequently, the current study excluded on-going depression and anxiety treatment information, to prevent making false conclusions about treatment effectiveness. A lack of effect of treatment on the subsequent course of depressive and anxiety symptoms in primary care patients has been previously shown (Penninx et al., 2011, Licht-Strunk et al., 2009b). As argued by Beekman et al. (2001) in many studies depression treatment data is available, but unused, due to the majority of older patients being untreated for this problem. The PROG-RES study did not include information on wider social network or the quality of interaction yet these variables may be important exploratory factors. In previous research, general social participation (Prince et al., 1998) and a large social network size (De Beurs et al., 2000) were significant predictors of a decreased likelihood of depressive disorders and anxiety symptoms respectively at follow-up in older adults.

6.6.4 Implications

Implications for clinical practice

The findings in chapter five and in this chapter focus attention towards three important clinical implications. One of the key finding reported in the previous chapter was establishing that people who consult primary with musculoskeletal pain who also have mild to severe depressive and anxiety symptoms at the initial presentation, 63% and 56% respectively, will have persistent symptoms over a 12-month period. Approximately 65% and 50% of individuals in these two groups had mild depressive and anxiety symptoms respectively, at each time point. Together with analyses of factors associated with cluster membership, this highlights the importance of enquiring about depressive and anxiety symptoms over time, including persistent mild forms that can also be associated with detrimental consequences for well-being.

Clearly a better understanding the course of depressive or anxiety symptoms in older people with musculoskeletal pain cannot be expected to result in immediate changes to clinical practice without strong evidence from other forms of research, including a need to improve current clinical practice. Nevertheless, assuming that the NHS would be interested in the recognition and management of persistent anxiety or depression symptoms in people with OA, this study draws attentions to the subject of how to identify patients with these problems. One possibility is annual evaluation of anxiety and depression symptoms severity, particularly among those older patients with more widespread pain, severe pain related disability, catastrophising or showing little interest in continuing meaningful activities regardless of OA.

To date, recognition and management of anxiety symptoms persistence

seems to be overlooked in health literature. Guidance for management of mixed anxiety and depression symptoms is also lacking- this has been regarded one of the main shortcomings of the current NICE depression guidance (Kendrick & Peveler, 2010). However, recent NICE depression guidance considered management of 'sub-threshold persistent depression symptoms', which is an important step towards the model of care reflecting cases typically seen in primary care. NICE (2009b) depression guidance considers persistent sub-threshold depression symptoms, defined as present for a considerable time, typically several months, despite 'active monitoring' or low- intensity treatment (for the DSM-IV and ICD-10 criteria of dysthymia, symptoms should be present for least 2 years) (Figure 1.2 on page18). Similar work is much needed for anxiety.

The following issue is what constitutes 'active monitoring'. It is currently undetailed in NICE guidance, but an interesting proposal for what it could be has been made by an academic general practitioner Lucassen et al. (2008). They outlined a stepped approach to diagnosing depression problem in general practice. This theoretical model implies that upon hearing the patient's story, the GP and the patient should agree on the name of the combination of symptoms and agree on the relative importance of the problem. Next, the role of a GP is to restore patient's beliefs in his/her own healing capacities. If despite of the forgoing activities the problem persists (there is no improvement in health), the use of medical diagnosis and treatment should be considered. Unfortunately this model remains only theoretical. Nevertheless, a recent systematic review of patient views on depression coexisting with a chronic physical illness (Alderson et al., 2012) seems to support this personalised approach to diagnostic labels, negotiating their meaning and management. The study found that patients hold different beliefs about

depression labels, its consequences, associated stigma, blame and responsibility and relevance of treatment, and patients' personal preferences may be important to how they engage in recognition of depression problem (Alderson et al., 2012). Whilst the study did not consider anxiety separately, one of the findings suggests that participants associated depression emotionally with anxiety (Alderson et al., 2012).

Another clinical implication of this study is an insight into factors potentially involved in adjustment to musculoskeletal pain, in the context of modern psychological theories. Interestingly, increased engagement in behavioural activities was found to have a positive effect on the course of depressive and anxiety symptoms. This finding supports a need for continuing enjoyable life activities by people experiencing musculoskeletal pain, a view advocated by the misdirected-problem solving and acceptance-commitment models. This stresses the importance of psychological interventions, which can modify unsuccessful coping strategies (e.g. catastrophising) and encourage more useful methods (e.g. increasing activities). Nevertheless, the persistence of anxiety and depression symptoms in older adults with musculoskeletal pain was found associated with a number of factors, including pain characteristics. Consequently, as advocated by NICE (2008) OA guidelines, successful adjustment to OA is likely to involve a multifaceted approach.

Research implications

Future research may consider modelling the dual trajectory of anxiety and depression symptoms inclusive of a limited number of time-varying covariates (e.g. number of pain sites or coping by increased behavioural activities), as this could result in additional clinical utility. It is also needed to assess contribution of factors

associated with the persistence of anxiety and depression symptoms coexisting with OA to an improved identification of these groups of primary care patients. This was previously advocated in the context of primary care patients in general (van den Brink et al., 2002).

Identification of primary care patients with musculoskeletal pain coexisting with persistent anxiety and depression symptoms is important, as they may benefit from specific anxiety and depression treatment. Research has demonstrated existence of potentially effective models of depression and anxiety care (Lin et al., 2003, 2006, Roy-Byrne et al., 2010). In practice, however, effective depression and anxiety care in older adults with musculoskeletal pain may be problematic. This issue is evident in the concerns raised by health professionals (Van Rijswijk et al., 2009, Barley et al., 2011). As such older primary care adults with depression are often undetected by their GPs (Licht-Strunk et al., 2009a) and older patients and patients with physical health problems often have their depression problem untreated (Kendrick et al., 2009). To date, little is known about this problem in older adults consulting primarily with musculoskeletal pain.

6.7 CONCLUSIONS

The study found that pain characteristics and coping strategies are the most prominent factors associated with the persistence of anxiety and depression symptoms in older people with musculoskeletal pain. Further research into the factors that may help to guide targeted identification of groups of individuals that would benefit most from depression or anxiety treatment is needed. In particular, future research may consider modelling the dual trajectory of anxiety and depression symptoms inclusive of a limited number of time-varying covariates, as

this could result in additional clinical utility. Understanding of factors associated with poor prognosis for depressive and anxiety symptoms have the potential to inform targeted case identification of patients with OA and coexisting depression or anxiety problems. However, to recognise a need for improvement, research is needed to understand the success of strategies implemented by GPs in primary care practice, to detect patients with possible or definite depression and anxiety problems.

Chapter seven: Documented detection of depression and anxiety in older adults consulting with musculoskeletal pain: analyses of medical record data

7.1 INTRODUCTION

Chapters three and five found that whilst formal depressive and anxiety disorders are relatively rare in people with OA/joint pain, persistent anxiety and depression symptoms are common in older adults consulting with musculoskeletal pain. The recognition, diagnosis and subsequent management of depressive and anxiety disorders in older patients and in particular those with chronic physical conditions, is challenging and consequently these conditions are likely to be undetected by general practitioners. To date, the detection of persistent anxiety and depression symptoms in older people with musculoskeletal pain is under-researched. This chapter considers the detection of depression or anxiety problems in general practice among older patients with musculoskeletal pain and coexisting persistent depression or anxiety symptoms (as identified by the LCGA anxiety and depression models). Patient medical records are reviewed to estimate the detection rate and factors associated with this rate are established.

7.2 BACKGROUND

7.2.1 Primary care challenges in detecting depression and anxiety

A critical challenge to detection of depression and anxiety in primary care practice is patient role in help-seeking. Typically it takes more than one consultation for the GP to detect mental health problems (Mitchell et al., 2009), with the low overall frequency of visits (combined with physical or pain presentation)

contributing to the further risk for the non-detection of depression in primary care adults (Menchetti et al., 2009). Underlying reasons for not seeking-help may be partially related to patient perception of the problem and the perceived need for health care (van Beljouw et al., 2010). A large study of primary care patients, with a confirmed diagnosis of depression and anxiety disorders, identified three distinct groups of patients who did not consult a health professional about their mental health problems. These three groups included patient's with no self-perceived problem; patient's with self-perceived problem but perceived no need for care and patient's with self-perceived problem and perceived unmet need for care (van Beljouw et al., 2010). The group with an unmet need for care had a comparably poor prognosis to those who consulted a health professional about depression or anxiety (van Beljouw et al., 2010). Some reasons underlying the unmet need for care were identified in a small qualitative study of primary care patients with depression or anxiety (Kadam et al., 2001). Some patients suggested that limited consultation time and apprehension to the use of pharmacological treatment (perceived to be the most commonly offered) prevent them from disclosing depression and anxiety (Kadam et al., 2001). Alternatively, patients may seek to meet their needs for depression and anxiety care outside of conventional health care (Kadam et al., 2001).

Identification of possible or definite depression and anxiety problems is important for formal diagnosis and subsequently the nature of intervention. Broadly speaking, general practitioners are suggested by clinical guidance to diagnose depression and anxiety, and tell their patients that they have a disease (NICE, 2005, NCC-MH, 2005, NICE, 2007, NICE, 2009b). Lucassen et al. (p.161, 2008) have controversially argued, that this way "*depression resembles appendicitis*". Primary care professionals report difficulties with the diagnosis of depression and anxiety in

older adults [Burroughs et al., 2006, Murray et al., 2006, Van Rijswijk et al., 2009] and in patients with chronic physical conditions [Van Rijswijk et al., 2009, Coventry et al., 2011]. In particular, they find problems when distinguishing 'true' cases from transient forms [Oladinni, 2002, Murray et al., 2006, Van Rijswijk et al., 2009, Coventry et al., 2011]. Recent research findings confirm this, with data from 32 sites participating in the Increasing Access to Psychological Therapies (IAPT) programme [Glover et al., 2010] showing that older patients (aged 65 and over) had fewer anxiety and depression diagnoses than those under 65. Some authors attribute this difficulty to somatisation that is common in older age and in people with chronic physical conditions [Tylee & Walters, 2007]. Furthermore, patients were often diagnosed with 'mixed anxiety and depressive disorders', highlighting the diagnostic problems arising from the frequent coexistence of these two conditions and the complex position of 'subclinical' conditions in the current classificatory systems [Glover et al., 2010]. Van Beljouw et al.'s (2010) study found that in a group of participants with a confirmed diagnosis of depressive or anxiety disorders and no record of contacts with a health care professional, those without self-perceived problems had the 'best course' for their depressive and anxiety symptoms (i.e. they recovered from anxiety and depression symptoms). This evidence supports the view that recognition may be less about categorising patients using the current classificatory systems, and more about detecting patients with self-perceived (potentially unexpressed) problems of depressive or anxiety symptoms and supporting them in making an informed decision about management [Lucassen et al., 2008, van Beljouw et al., 2010]. Although, clearly this view may be difficult to apply to patients who potentially fail to acknowledge their depression/anxiety problems, due to underlying psychological reasons, such as

fear of stigma - a problem commonly highlighted for depression by patients themselves [Alderson et al., 2012].

Small qualitative studies suggest that the availability of access to relevant interventions may affect primary care professional willingness to identify mental health problems [Burroughs et al., 2006] and for patient willingness to disclose them [Kadam et al., 2001]. Studies consistently show that treatments for depression and anxiety are under-utilised in older adults. For example, a large dataset of medical records of UK primary care patients with depression symptoms found that older patients and patients with chronic physical problems were less likely to be treated for depression [Kendrick et al., 2009]. Similarly, in a large randomised controlled trial of older primary care patients with major depression or dysthymia (of whom approximately half had arthritis [Lin et al., 2003, 2006]) only 65% reported any lifetime depression treatment [Unützer et al., 2003]. A systematic review of anxiety in older adults suggests that anxiety treatment utilisation may also be lower in older adults than in younger adults [Wolitzky-Taylor et al., 2010]. Prescribing pharmacological treatment is a core professional skill of GPs yet this type of treatment may be poorly utilised in older patients with musculoskeletal pain, as these patients are known to be highly vulnerable to the adverse effects of psychotropic medications [Bulat et al., 2005], for example, benzodiazepines have been shown to increase the risk for falls and skeletal fractures [Bulat et al., 2008]. Non-pharmacological treatment may also be poorly utilised in older patients with musculoskeletal pain, as primary care practitioners consistently expressed a sense of poor provision or problems with access to specialist mental health, volunteer or social services [Rogers et al., 2001, Telford et al., 2002, Pollock & Grime, 2003, Burroughs et al., 2006].

7.2.2 Is it important to detect depression and anxiety in older people with musculoskeletal pain?

Regardless of the difficulty associated with detecting older primary care patients with musculoskeletal pain and coexisting problems of depression and anxiety, there are reasons to strive for detection, which goes beyond simply reducing the unpleasant symptoms associated with depression and anxiety. As discussed in chapter one, there is robust evidence to suggest the adverse effects of anxiety and depression symptoms can result in poorer OA-related outcomes. The Improving Mood-Promoting Access to Collaborative Treatment (IMPACT) trial offers promising potential improving OA-related outcomes through detection of possible or definite depression. This study included 1001 older primary care attendees with self-reported arthritis (mostly OA) and clinically diagnosed major depression and/or dysthymia and compared usual care (i.e. primary care treatment and potential referral to specialty mental health care) with a collaborative care approach (i.e. a case manager collaborating with a patient and primary care providers). The latter was not only associated with fewer depression symptoms, but importantly also improved OA-related outcomes, such as pain, functional outcomes and improved quality of life (Lin et al., 2003, 2006). The inclusion of patients with dysthymia, suggest that the result can be generalised to patients with persistent, subclinical symptoms of depression.

A large RCT that was modelled on the IMPACT trial offers preliminary insight into the impact of detection of anxiety disorders in general primary care patients (Roy-Byrne et al., 2010). The study excluded milder forms of anxiety disorders and unstable medical conditions. Nevertheless, relative to usual care, collaborative care led to a greater reduction of in anxiety and depression symptoms and

improved functional status. An understanding of how successful this model is in people with OA warrants a similar trial in this sub-population of primary care patients.

Clearly, it can be argued that collaborative depression and anxiety care may be unnecessary, if there is ample evidence to demonstrate that OA treatment can lead to clinically significant improvements in concurrent depression and/or anxiety. However, systematic reviews of the main modalities of OA pain control, including exercise (Fransen et al., 2009, Brosseu et al., 2010), paracetamol (Towheed et al., 2009), anti-inflammatory medication (Watson et al., 2006, Derry et al., 2012), opioids (Nüesch et al., 2010) seldom summarise the effects on coexistent depression and/or anxiety. For example, conclusions of a systematic review of exercise, including two studies (in this one review of five studies) were limited to a statement that participants reported "*decreased depression and anxiety*" (Brosseu et al., 2010).

7.3 RATIONALE OF THE STUDY

Despite effective treatment options for anxiety (Wolitzky-Taylor et al., 2010) and depression (Yohannes & Caton, 2010), detection of these conditions in primary care is challenging. According to Timonen and Liukkonen (2008), cross-sectional studies suggest that between 50% and 70% of patients with depression in primary care settings are undetected. Based on a 3-year follow-up period, 37% of primary care patients with depression or anxiety were not detected (Kessler et al., 2002). A much higher rate - 67% - was reported for older primary care patients with depressive disorders followed-up for one year (Licht-Strunk et al., 2009a). Exact estimates of detection rates may be affected by the methodology used (i.e. symptoms persistence (Licht-Strunk et al., 2009a), period of detection (Kendrick, 2008), definition of detection (Joling et al., 2011)).

A primary care study of 250 adult patients found that the odds of non-detection of depression were higher for patients presenting with physical symptoms (OR: 2.3; 95% CI 1.1, 5.3) and specifically with pain (OR: 4.1; 95% CI 1.6, 9.9) (Menchetti et al., 2009). Overall, older adults with coexisting musculoskeletal pain and clinically significant depression symptoms seem to be at a high risk of being under-detected in primary care. Much less is known about anxiety, but given that anxiety in general, received limited research attention in people with musculoskeletal pain; detection of anxiety is also likely to be poor.

7.4 AIM AND OBJECTIVES

The **aim** of this study is to describe the detection of depression and anxiety in older primary care patients with OA.

Specific objectives are:

- To determine the detection rate of persistent depression or anxiety symptoms in older patients presenting to general practice with musculoskeletal pain
- To establish the factors associated with detection in this sub-population

7.5 METHOD

7.5.1 Selection of sample

PROG-RES participants were eligible for this analysis if they consented for medical record review and were classified in the refined LCGA models (see section 6.4.2 on page 205) as having *persistent anxiety symptom or depression symptom trajectories*. A practical decision of focusing on patients with either depression or anxiety was made to increase the statistical power of analyses and

to account for an overlap in range of treatments used to manage depression and anxiety in the community. Focusing on persistent symptoms was grounded in a suggestion that relative to cross-sectional estimates, focusing on persistent symptoms might be more relevant for patient care by identifying more enduring and potentially more clinically relevant symptoms (Licht-Strunk et al., 2009a, Luccasen et al., 2008).

7.5.2 Data selection

Evidence of detection

Data indicative of the documented *detection* of depression and anxiety were selected for extraction from the medical records. This involved using a definition similar to those that have been used in previous studies (e.g. Licht-Strunk et al., 2009a, Joling et al., 2011, Cully et al., 2009).

The first group of indicators included a Read-coded diagnosis of depression and/or anxiety symptoms (Licht-Strunk et al., 2009a, Joling et al., 2011, Cully et al., 2009). Henceforth this group of indicative medical records will be called documented *diagnosis/problem codes*.

However, reliance only on recorded diagnosis is likely to result in an underestimation of *detection* rates for depression and anxiety (Joling et al., 2011). Records of antidepressant prescription were found to be the best single predictor of GP's *detection* of depression, assessed with the Composite International Diagnostic Interview (CIDI) (Joling et al., 2011). As a result, and in line with previous research (Kendrick et al., 2009, Licht-Strunk et al., 2009a, Cully et al., 2009), evidence of the provision of treatment for mental health problems was also extracted. This included receiving antidepressants and anxiolytic drug prescriptions in addition to referral to

mental health or appropriate social services. This group will be referred to as documented *interventions*. In summary, for the purpose of the analysis presented in this chapter documented *detection* encompasses evidence from the electronic medical record (EMR) of documented anxiety- and depression-related *diagnosis/problem codes or interventions*.

Anxiety and depression-related Read-codes were identified using a list of clinical search terms and Read codes (5-byte, Version 2). These were generated using the NHS Information Authority ((NHSIA), 2000) Clinical Terms Version 3. A provisional list of search terms was enhanced by comparison with lists of search terms previously reported in studies into the diagnosis of depression (Rait et al., 2009) and antidepressants use (Coupland et al., 2011). Whilst medical records data for anxiety diagnosis and intervention has also been reported (Goodman & Tyer-Viola, 2010, Stein et al., 2004), lists of search terms used in the identified studies were unavailable. The provisional list of search terms was discussed with an academic general practitioner (CM). For the agreed list of search terms see Appendix F.1 (p. 391).

Records reflecting the diagnosis of depression and anxiety included broad (*feeling stressed*, *emotional upset*, *low mood*, *anxiousness*) and narrow terms (e.g. *[X] Depressive episode*, *[X] Anxiety state*). A term *postnatal depression* was excluded, as this condition was unlikely to be related to older general practice attendees.

Medication used to identify treated cases of depression and anxiety included following classes of antidepressants and anxiolytic drugs (as described in the British National Formulary (BNF) 61 (Joint Formulary Committee, 2011): depression-specific medication (e.g. tricyclic and related antidepressants, monoamine oxidase inhibitors, other antidepressants, compound antidepressants); anxiety-specific

medication (e.g. benzodiazepine or its derivative drugs and propranolol) and drugs that are indicated for both depression and anxiety (e.g. selective serotonin reuptake inhibitors, selective-norepinephrine reuptake inhibitors). Medical records were examined in detail, to ensure that medications were actually prescribed for depression and anxiety. Following a study in patients with heart failure, inclusion of trazodone and amitriptyline, was determined by excluding possibility of their use for other health problem, such as sleep and pain [Cully et al., 2009]. The BNF 61 [Joint Formulary Committee, 2011] was used as a reference point. Although in general the BNF does not recommend amitriptyline for depression, 30 mg - 75 mg is suggested for depression, if needed gradually increasing to 150 mg - 200 mg. For neuropathic pain/migraine prophylaxis, the BNF suggests 10 mg, if necessary followed by a gradual increase to 75 mg [Joint Formulary Committee, 2011]. The BNF recommends 75 mg to 300 mg trazodone for depression/anxiety [Joint Formulary Committee, 2011]. Exclusions were informed by discussions with an academic general practitioner (CM).

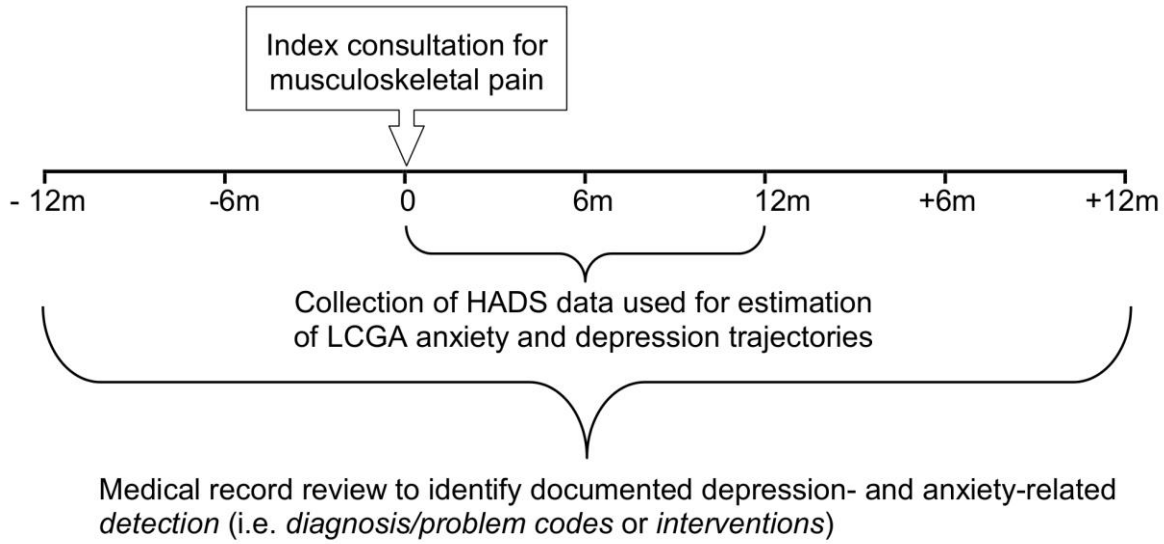
Psychological interventions were identified using methods described in the existing literature [Kendrick et al., 2009] (e.g. referral to counselors, psychologists, mental health or social workers, psychiatry services and occupational therapists). Any mental health encounters or clinic visits [Cully et al., 2009] or psycho-education or management plan (e.g. advice and crisis plans) were included.

Time frame for extracting data

As per previous research [Licht-Strunk et al., 2009a, Cully et al., 2009] analyses focused on a one year period following the index consultation, as the anxiety and depression trajectories were based on HADS data collected in that time period. In addition, the period of medical record review was extended one year either side of

the period of observation. A timeline of medical record reviews conducted for the selected sample can be seen displayed in Figure 7.1.

Figure 7.1 Timeline of data extracted from medical records of PROG-RES participants eligible for the current study.



In addition to exploring changes in the frequencies of *detection* over time, the choice of the three year period was done for two main reasons. Firstly, the analyses aimed to acknowledge the 'real' circumstances of making a clinical diagnosis. The likelihood of identification of psychological problems in primary care patients increases with the number of consultations (Roy-Byrne et al., 2000, Mitchell et al., 2009). This could be related to an increasing number of opportunities to prompt recognition or implementation of 'watchful waiting' by GPs (Simon & Von Korff, 1995), a method advocated by NICE (2009a, 2009b). Secondly, it was unlikely that the onset of anxiety and depression symptoms always coincided with the start of the study. As reported in chapter four, participants were likely to have disabling pain at the index consultation, and thus, the process of emotional adjustment to pain had

been initiated before the study commenced. In addition, numbers of new episodes of depression and anxiety at follow-up were low, as indicated by HADS scores.

Data extraction procedures

Data on all recorded consultations was extracted from the electronic medical records (EMR) for all patients consenting to review (n=428) by a data custodian. The data custodian provided dates of the index consultation and based on the pre-defined list of Read terms and Read codes extracted relevant consultations, referrals and prescription information from the EMR for the 428 participants. During data extraction, both the data custodian and the principal investigator (MR) were blinded to the anxiety and depression trajectories of participants. Raw data were presented in the PASW long format. The index consultation dates were presented along with participants' anonymised IDs. Raw data for consultations, referrals and prescriptions were provided separately and included: participants' IDs, corresponding Read terms and dates of recording. Consultation and referral data also included specific Read codes and prescription data included relevant BNF chapters.

7.5.3 Statistical analyses

Rates of detection

PASW Statistics version 20 was used for analyses of *detection* rates. Simple frequencies of occurrence of each depression- and/or anxiety- Read terms grouped as *diagnosis/problem codes* or *interventions*, were estimated using the entire 3-year period of record review. In cases when the same Read term was used more than once for the same patient it was counted as one record. The same

medications with different dose or in different forms were regarded to be the same Read term.

The overall proportion of eligible participants with any documented evidence of *diagnosis/problem codes* and *interventions* were calculated separately and were not mutually exclusive. The overall frequency of all eligible participants with either Read coded diagnosis/problem codes or interventions, was calculated to form a basic estimate of the rate of *detection* of anxiety or depression over the 3-year period of medical records review. The total number of times a relevant evidence of *detection* (i.e. *diagnosis/problem codes* or *interventions*) was also recorded in each 6-month interval.

Factors associated with detection

The numbers of all consultations were recorded for eligible participants by estimating time between the index consultation and identified anxiety- or depression-related Read codes. Average anxiety and depressive scores for baseline, 3-, 6- and 12-month follow-up scores were calculated. For consistency with the analyses described in chapter five, the following baseline covariates were also included: age (50-59, 60-69, 70 years or more); gender (male/female); living alone (yes/no); lack of partner (yes/no); manual/routine occupational class (yes/no); emotional and instrumental support (yes/no); coping (split by the highest tertile) by catastrophising, ignoring pain, coping-self statements, and increased behavioural activities; widespread pain (number of pain sites). Due to the issue of multicoolinerity reported in section 6.5.2 (p. 208) for variable ‘pain interference with daily activities’, the average of pain interference (0-10 NRS) with daily, social and work activities was taken. This was justified as these are typically combined in the

Chronic Pain Grade (see section 4.4.5, p. 148). The median was used to describe all continuous variables, due to non-normal distribution of scores (see Table 4.3 (p. 146) for HADS scores, Appendix D.1 (p. 374-376) for pain interference with activities scores and the numbers of pain sites, Appendix F.2 (p. 395) for the numbers of all consultations).

Statistical analyses were conducted in three parts, using PASW statistics version 20 in the first two parts and the STATA software (version 11.1) in the last part. In the first part, the differences between detected and undetected patients were determined across the selected factors. Due to non-normal distribution of scores, the differences across average anxiety and depressive HADS scores, widespread pain and interference with activities were determined using the Mann-Whitney U test. This test was selected as data failed to meet the assumptions required for parametric tests. Differences in categorical data were compared using the Chi^2 test or for expected cell sizes of less than 5, the Fisher's exact test, for which statistical significance was established at the level of $p<0.05$.

In the second part of the analyses, due to non-normal distribution of scores, median numbers of consultations were estimated across categorical data and correlations between continuous data and numbers of all consultations were estimated using the Spearman Rank Correlation test. This aimed to assess interactions between the frequency of consulting and other factors assessed for their associations with *detection* of depression or anxiety.

In the third part, a multivariable logistic regression analysis was conducted for variables that yielded $p<0.10$ in the first part of analyses. Borderline variables were then included and the model was re-estimated, to assess changes in statistical significance (established at the level of $p<0.05$).

7.6 RESULTS

7.6.1 Sample size

In total, 143 participants were eligible for this analysis, as they consented for medical record review and were classified in the refined LCGA models (section 6.5.1) as having persistent anxiety or depression symptoms.

7.6.2 General descriptive findings

Read terms of diagnosis/problem codes and interventions that were identified in the medical records of PROG-RES participants with persistent depression or anxiety trajectories are displayed in Tables 7.1 overleaf and 7.2 on page 244 respectively. Overall frequencies were not provided, as multiple consultation terms were used for some participants. Read terms *anxiety with depression* and the medication *diazepam* were the most commonly used diagnosis/problem codes and interventions Read terms respectively.

Table 7.1 *Diagnosis/problem codes* Read terms and frequency of their occurrence for the 143 eligible participants.

Read terms	Frequency of occurrence‡ (%)	
Anxiety with depression	11	(7.7)
Low mood	7	(4.9)
Depressed	5	(3.5)
Depression	4	(2.8)
Stress related problem	4	(2.8)
[X]Depressive episode	3	(2.1)
Panic attack	3	(2.1)
[X]Anxiety NOS	2	(1.4)
[X]Anxiety reaction	2	(1.4)
[X]Depression NOS	2	(1.4)
Recurrent depression	2	(1.4)
[X]Mild depressive episode	1	(0.7)
[X]Recurrent depressive disorder	1	(0.7)
Acute reaction to stress	1	(0.7)
Anxiety states	1	(0.7)
Chronic anxiety	1	(0.7)
Emotional upset	1	(0.7)
Panic disorder	1	(0.7)
Pt health quest (PHQ-9) score	1	(0.7)
Reactive Depression	1	(0.7)

Note: NOS- not otherwise specified; [X] - ICD-10 criteria based;
‡- If the code was used more than once for the same person, it was counted as used once.

Table 7.2 *Interventions* Read terms and frequency of their occurrence for the 143 eligible participants.

Read terms		Frequency of occurrence ‡ (%)	
Diazepam	(tablets 2 mg, 5 mg)	25	(17.5)
Citalopram	(hydrobromide tablets 10 mg, 20 mg, 40 mg)	20	(14.0)
Sertraline	(hydrochloride tablets 50 mg, 100 mg)	8	(5.6)
Escitalopram	(tablets 5 mg, 10 mg, 20 mg)	7	(4.9)
Dosulepin	(capsules 25mg, tablets 75 mg)	5	(3.5)
Venlafaxine	(hydrochloride m/r capsules (75 mg, 150 mg), hydrochloride tablets 37.5 mg)	5	(3.5)
Paroxetine	(hydrochloride tablets (20 mg, 30 mg), tablet 20 mg)	4	(2.8)
Refer to counsellor		4	(2.8)
Propranolol	(hydrochloride tablets 40 mg, hydrochloride m/r capsules 80 mg)	3	(2.1)
Fluoxetine	(hydrochloride capsules 20 mg)	2	(1.4)
Mirtazapine	(orodispersible tablets (15 mg, 30 mg), tablets 15 mg)	2	(1.4)
Lofepramine	(tablets 70 mg)	1	(0.7)
Lorazepam	(tablets 1 mg)	1	(0.7)
Lormetazepam	(tablets mg)	1	(0.7)
Nortriptyline	(tablets 10 mg)	1	(0.7)
Refer to occupational therapist		1	(0.7)
Seen in psychiatric unit		1	(0.7)
Trimipramine	(tablets 25 mg)	1	(0.7)

Note: ‡- If the same medication (regardless of dose) was used more than once for the same person, it was counted as used once.

7.6.3 Documented depression and anxiety related *detection*

Table 7.3 Frequencies of documented *detection* (i.e. *diagnosis/problem codes* or *interventions*) in older primary care patients with musculoskeletal pain†.

		Persistent anxiety/depression symptom trajectories					
		Anxiety or depression combined	**Overlapping groups**		**Mutually exclusive groups ‡**		
			Anxiety	**Depression**	**Anxiety only**	**Depression only**	**Anxiety and depression**
		(n = 143)	**(n =113)**	**(n = 87)**	**(n = 55)**	**(n = 27)**	**(n = 57)**
Detection	n (%)	61 (42.7)	51 (46.0)	37 (42.5)	23 (41.8)	8 (29.6)	28 (49.1)
Diagnosis/problem codes	n (%)	37 (25.9)	34 (30.1)	24 (27.6)	13 (23.6)	3 (11.1)	21 (36.8)
Interventions	n (%)	57 (39.9)	48 (42.8)	36 (41.4)	20 (36.4)	8 (29.6)	27 (47.4)

Note: †- frequencies of *detection* are not equal to the total number of *diagnosis/problem codes* or *interventions* as some participants were identified by more than one method; ‡- in four cases depression or anxiety trajectories only were available, and so mutual exclusiveness could not be judged.

Table 7.3 (on the previous page) shows that having persistent HADS defined depression symptoms only was associated with a lowered rate of *detection* (29.6%). Of the 143 patients with persistent anxiety or depression symptoms over the 12-month period of observation, 42 (29.4%) had evidence of *detection* from the general practice medical record during the same 12-month period of having their depressive and anxiety symptoms assessed with the HADS. The above figure, however, may under-estimate the true rate of *detection* since extending the period of medical record review one year either side of the period of observation identified a further 19 detected cases giving a total estimated *detection* rate of 42.7% over 3 years (Table 7.4). In fact, the majority (n=35) of detected cases had already been identified before the index consultation (Table 7.4). Evidence of *detection* came primarily from documented *interventions*, with just 4 out of 61 cases being detected on the basis of *diagnosis/problem codes* alone (Table 7.4).

Table 7.4 Frequencies of documented depression- or anxiety-related *diagnosis/ problem codes*, *interventions* and overall *detection* over time for the 143 eligible participants.

Time point	*Diagnosis/ problem codes*		*Interventions*		*Detection*	
	N (%)	**N (cum %)**	**N (%)**	**N (cum %)**	**N (%)**	**N (cum %)**
- 12m	11 (7.7)	11 (7.7)	25 (17.5)	25 (17.5)	30 (21.0)	30 (21.0)
- 6m	7 (4.9)	14 (9.8)	27 (18.9)	32 (22.4)	28 (19.6)	35 (24.5)
6m	9 (6.3)	19 (13.3)	29 (20.3)	41 (28.7)	31 (21.7)	45 (31.5)
12m	17 (11.9)	28 (19.6)	31 (21.7)	50 (35.0)	35 (24.5)	53 (37.1)
+ 6m	14 (9.8)	32 (23.4)	31 (21.7)	54 (37.8)	35 (24.5)	57 (39.9)
+12m	13 (9.1)	37 (25.9)	34 (23.8)	57 (39.9)	36 (25.2)	61 (42.7)

7.6.4 Patient characteristics associated with documented *detection* of depression or anxiety

Table 7.5 overleaf presents the results of the characteristics of patients with detected vs. undetected anxiety and depression symptoms during the three year data extraction period. The result of the Mann-Whitney U test suggests that the total number of consultations was most strongly associated with *detection*. In other words, the 61 detected patients had significantly more opportunities for identification of their mental health symptoms than the 82 undetected patients as they consulted their doctor more frequently. The result of the Mann-Whitney U test indicates that the detected patients had significantly higher aggregated HADS-anxiety scores, with the effect less robust than for numbers of consultations. According to the Man-Whitney U tests (found in Table 7.5 overleaf), patients who had their depression detected by the general practitioner had significantly higher pain interference with activities (daily activities, social activities and work activities). Women were more likely to have their depression or anxiety detected than men, with a 26% (95% CI 11.3, 41.2) difference in the proportion of females across the two groups. Table 7.6 on page 250 shows that higher levels of *detection* in women and in people with higher pain interference with different activities could be associated with these individuals being more likely to consult in general. Patients with more severe depression symptoms appeared also to have higher rates of *detection* although this finding was of borderline statistical significance. A multivariable logistic regression analysis (for variables that met the entry probability $p<0.10$ for observed significance level) was conducted. Decreasing number of consultations (OR: 0.98 (95% CI 0.96, 0.99), $p<0.0001$) and average HADS-anxiety score (OR: 0.82 (95% CI 0.70, 0.96), $p=0.012$) were found to be significantly associated with the risk of being undetected (Table 7.5 overleaf).

Table 7.5 Characteristics of older primary care patients with musculoskeletal pain, detected vs. undetected by the GP during the three year data extraction period.

Variable			Detected (n=61) n (%)		Undetected (n=82) n (%)		Observed significance level	Multivariable logistic¥ regression with entry probability p<0.10
Number of consultations		Med. (IQR)	78	(61.0)	46	(32.5)	MED 31.0 (95% CI 19.0, 43.0), p<0.0001	OR 0.98 (95% CI 0.96, 0.99), p<0.0001
Averaged HADS-A†		Med. (IQR)	10.7	(3.8)	9.5	(3.3)	MED 1.25 (95% CI 0.50, 2.25), p= 0.003	OR 0.82 (95% CI 0.70, 0.96), p=0.012
Averaged HADS-D†		Med. (IQR)	8.3	(4.4)	7.8	(4.0)	MED 0.92 (95% CI -0.25, 2.00), p=0.115*	-
Age	50-59		23	(37.7)	31	(37.8)	p=0.598	-
	60-69		19	(31.1)	20	(24.4)		
	70+		19	(31.1)	31	(37.8)		
Gender	Males		13	(21.3)	39	(47.6)	p=0.001	OR 0.49 (95% CI 0.21, 1.15), p=0.103
	Females		48	(78.7)	43	(52.4)		
Manual/ routine work								
	No		39	(63.9)	50	(61.0)	p=0.718	-
	Yes		22	(36.1)	32	(39.0)		
Living alone								
	No		48	(78.7)	65	(79.3)	p=0.933	-
	Yes		13	(21.3)	17	(20.7)		
Partner	No		17	(27.9)	20	(24.4)	p=0.597	-
	Yes		43	(70.5)	62	(75.6)		
Emotional support								
	No		7	(11.5)	4	(4.9)	p=0.203	-
	Yes/no need		51	(86.9)	78	(95.1)		
Instrumental support								
	No		11	(18.0)	8	(9.8)	p=0.149	-
	Yes/no need		50	(82.0)	74	(90.2)		

Table 7.5 cont. Characteristics of older primary care patients with musculoskeletal pain, detected vs. undetected by the GP during the three year data extraction period.

Variable	Detected (n=61) n (%)	Undetected (n=82) n (%)	Observed significance level	Multivariable logistic¥ regression with entry probability p<0.10
Catastrophising				
No	25 (41.0)	45 (54.9)	p=0.100*	-
Yes	30 (49.2)	30 (36.6)		
Ignoring pain				
No	40 (65.6)	51 (62.2)	p=0.165	-
Yes	11 (18.0)	25 (30.5)		
Coping-self statements				
No	34 (55.7)	44 (53.7)	p=0.590	-
Yes	21 (34.4)	33 (40.2)		
Increased behavioural activities				
No	27 (44.3)	41 (50.0)	p=0.737	-
Yes	26 (42.6)	35 (42.7)		
Number of pain sites Med. (IQR)	11 (17.0)	9.0 (11.0)	MED 2.0 (95% CI -1.0, 5.0), p=0.146	-
Interference with activities Med. (IQR)	7.3 (4.0)	6.0 (3.3)	MED 1.0 (95% CI 0.0, 2.0), p=0.024	OR 0.93 (95%CI 0.78, 1.10), p=0.413

Note: IQR- inter quartile range; Med.- median; MED- median difference;
Some proportions are not equal 100% due to missing values;
†- averaged baseline 3, 6 and 12 months scores;
¥- The model was statistically significant (LL= -73.19, LR chi^2= 37.48, Prob > chi p<0.001, Pseudo R^2= 0.21);
* - The impacts of borderline variables 'average HADS-D score' and 'catastrophising' were investigated, with 'number of consultations' and 'averaged HADS-A' remaining to be the only statistically significant variables.

Table 7.6 Number of consultations across different characteristics of the 143 older primary care patients with musculoskeletal pain.

Variable	Median numbers of consultations (IQR)
Averaged HADS-A† Spearman rho	0.047, p= 0.577
Averaged HADS-D† Spearman rho	0.157, p= 0.061
Age	
50-59	45.5 (41.5)
60-69	66.0 (50.0)
70+	65.5 (53.0)
Gender	
Males	44.5 (38.5)
Females	65.0 (55.0)
Manual/routine work	
No	58.0 (51.5)
Yes	57.5 (52.5)
Living alone	
No	54.0 (50.0)
Yes	65.5 (58.0)
Partner	
No	73.0 (51.5)
Yes	54.0 (44.0)
Emotional support	
No	57.0 (72.0)
Yes/no need	58.0 (51.0)
Instrumental support	
No	58.0 (52.0)
Yes/no need	58.0 (53.0)
Catastrophising	
No	51.0 (45.8)
Yes	63.0 (63.5)
Ignoring pain	
No	58.0 (50.0)
Yes	50.5 (50.5)
Coping-self statements	
No	61.5 (49.0)
Yes	53.0 (53.8)
Increased behavioural activities	
No	57.5 (40.0)
Yes	56.0 (58.5)
Number of pain sites Spearman rho	0.066, p=0.431
Interference with: Spearman rho	
Daily activities	0.217, p=0.012
Social activities	0.176, p=0.040
With work	0.237, p=0.005

Note: IQR- inter quartile range; †- averaged baseline, 3, 6 and 12 months scores.

7.7 DISCUSSION

7.7.1 Summary of main findings

This study suggests that among older patients consulting general practice with musculoskeletal pain who also have coexisting persistent anxiety or

depression symptoms, just under half will have documented evidence that their mental health problems have been detected by the GP. The highest *detection* rate was found in patients with both persistent depressive and anxiety symptoms (49%). Patients with a higher numbers of consultations and those with more severe anxiety symptoms (adjusted for gender and pain inference with activities) were significantly more likely to be detected. Females and those with higher levels of pain that interfered with activities appeared more likely to have their coexisting depressive or anxiety symptoms identified (although these patients were, in general, more likely to have consulted their GPs). Patients with more severe depression symptoms and often catastrophising appeared in unadjusted analyses to have higher rates of *detection*, but this association was of borderline statistical significance.

7.7.2 Comparison with previous findings

Direct comparisons between this study and other studies, reporting detection of anxiety and depression in general practice are problematic for three main reasons. Assessing true differences is challenging in the light of varying methodologies, especially the period of observation, and the extracted evidence or independent consideration of depressive and anxiety symptoms. Identifying diagnosed depressive or anxiety disorders in primary care patients in general (the focus of many previous studies) may not represent a comparable challenge to detecting persistent, but largely mild to moderate symptoms among older patients with musculoskeletal pain. Analyses of factors associated with anxiety and depression symptom detection often have limited statistical power as they are often based on small samples – a problem acknowledged in the context of

depression detection by Licht-Strunk et al. (2009a). Nevertheless, with attention to these challenges, several comparisons are possible.

Detection rates

In Kessler et al.'s (2002) study, 88 primary care patients with clinically diagnosed depressive and anxiety disorders were followed for three years. In total, 56 (67%) patients had medical record evidence indicative of detection (i.e. psychological diagnoses, treatments, and referrals). Given the methodological comparability to this thesis, 67% detection rate in general primary care patients, seemed greater than 43% found in the study population investigated in this thesis. In contrast, after taking into consideration methodological differences, 43% appeared broadly comparable with 33% (Licht-Strunk et al., 2009a) and 52% (Volkers et al., 2004) detection rates, reported for older primary care patients with depressive disorders. Overall, it is likely that older primary care consulters with musculoskeletal pain are at a high risk of non-detection of their depression or anxiety problem.

Factors associated with detection

The factors associated with *detection* are similar to those reported in previous studies. As originally reported by Menchetti et al. (2011), this study also found that the number of consultations are positively associated with detection rates. As consultation rates in England in 2007/2008 tended to be higher for females than for males (Hippisley-Cox & Jumbo, 2008), in this thesis women were found to have higher consultation rates and as such were more likely to have their anxiety and depression problems detected.

As reported by other studies, patients with more severe anxiety symptoms were found to be more likely to receive mental health treatment (Weisberg et al., 2007, van

Beljouw et al., 2010, Licht-Strunk et al., 2009a). However, in contrast with previous research, the severity of depression symptoms was not significantly associated with detection (Dowrick & Buchan, 1995, van Beljouw et al., 2010, Licht-Strunk et al., 2009a). In the current sample having depression symptoms only was much less frequent (7% of 502 respondents to baseline questionnaire) than anxiety symptoms only (23%) (see Table 4.2 on page 147). This could contribute to the observed lack of significant effect of depression symptoms severity on the likelihood of being detected.

7.7.3 Strength and critical considerations

Strength

This study has two major strengths. It addresses some methodological issues previously identified as important to investigations of detection rates. Namely, numbers of opportunities for detection (Roy-Byrne et al. 2000, Mitchell et al., 2009), the longitudinal nature of clinical decision making (Dowrick & Buchan, 1995, Licht-Strunk et al., 2009a) and the choice of a variety of indicators when determining whether comorbid symptoms have been identified (Joling et al., 2011). This was achieved by using three years of medical record data and by selecting records of both *diagnosis/problem codes* (formal Read code in the medical records) and *interventions* (a range of treatments and referrals) for depression and anxiety.

To the best of the author's knowledge, this is the first study focussing on patients with persistent mental health symptoms. Given that mild to moderate depressive and anxiety symptoms are more common than formally diagnosed depressive and anxiety disorders in people with OA (chapter two), the use of repeated measures to assess anxiety and depression symptoms is a major strength when looking at the wider issue of persistent depressive and anxiety symptoms.

Critical considerations

Factors that could influence the behaviour of GPs participating in the PROG-RES study require critical considerations. PROG-RES procedures involved GPs screening for depression symptoms at the index consultation with two questions. It could be argued that this could inflate the estimated *detection* rate. A comparison of HADS scores and the results of GP-administered screening, with a large number of patients with depression symptoms being missed, suggests that screening was unlikely to improve identification of depression symptoms in PROG-RES participants (Mallen & Peat, 2008). Still, the results of screening could be influenced by the GP's awareness of the participant's medical history. Nevertheless, as found in this study, the majority of cases were detected before the index consultation and no marked increase in numbers of detected patients could be observed within six months from the index consultation (see Table 7.4 on page 246).

Another factor that could impact GPs' behaviour is the availability of clinical recommendations during the reviewed period of data extraction. From April 2006 the UK general practice contract has incentivised general practitioners to systematically assess for the presence of depression in patients with heart disease and diabetes using standardised assessment tools. National clinical guidelines for depression and anxiety were introduced in 2004 and 2007 respectively, guidance for depression in chronic illness in 2009. PROG-RES participants were recruited between September 2006 and March 2007, and thus, the current study included records documented between September 2005 and March 2009. It is unclear if the national guidance affected decision making of GPs participating in the PROG-RES. Assuming that guidance for depression in a chronic physical illness in 2009 might affect recognition and management of this problem in OA, studies based on

medical records reviewed after year 2009 could result in a higher estimate than in the current study.

A further factor that could influence the results of this study relates to the individual practices for GPs coding of consultations. The actual detection rate could be higher if GPs are failing to code consultations, rather than representing a lack of awareness (Ani et al., 2008, Joling et al., 2011). It is not possible to estimate the exact scope of non-recording in the current study. However, a prospective uncontrolled intervention study was conducted in the Keele General Practice Research Partnership research network, with coding found to be variable across practices, but repeated assessments, feedback, and training appeared to improve data quality of coding (Porcheret et al., 2004). Practices used in the PROG-RES are a part of the Keele General Practice Research Partnership, which conduct regular audits of their coding practices, so they are of high quality Read coding (Porcheret et al., 2004).

This thesis excluded free-text recording in the electronic medical records, which include additional information (e.g. from diagnostic tests, operatives' reports and consultation letters (Tu et al., 2010)). Tu et al. (2010) have argued that manual tagging of free-text data is heavily reliant upon personal judgment, which can lead to misinterpretations (Tu et al., 2010). Joling et al. (2011), in their study of detection rates of depression in general practice, used not only medical records of psychological diagnosis (including both depression and anxiety), treatment and referrals, but also added free-text records. They reported 69% of the overall detection rate - a number that appears broadly comparable to 63% reported by Kessler et al. (2002) for depression and anxiety (obtained without searching free-text). Together there are no strong reasons to believe that detection rates would be altered if free-text search had been undertaken in the current study. Additionally, the estimates provided in this study can be regarded reliable, as a

comprehensive list of indicators of detection was used, which are comparable to indicators used in previous research (Kessler et al., 2002, Kendrick et al., 2009, Licht-Strunk et al., 2009a, Cully et al., 2009).

7.7.4 Implications

Implications for clinical practice

Assuming that the detection of depressive and anxiety problems in people with OA is of interest to the NHS, targeted case identification (using screening tools) has been proposed as a method of early detection. This method has been suggested in the context of managing the low levels of detection of depression in older patients who are at high risk (Licht-Strunk et al., 2009a) and in patients with painful conditions (Menchetti et al., 2009) and adults with anxiety in primary care (Stein et al., 2005). Yet, three systematic reviews in adult primary care attendees found that when used in isolation, routinely administered stand-alone screening questionnaires to assess anxiety or depression have little impact not only on detection, but also intervention used and clinical outcome (Pignone et al., 2002, Gilbody et al., 2005, 2008).

Whilst the efforts of care providers to implement high quality mental health care should not be underestimated (Stein et al., 2005), Cully et al. (2009) highlight that the identification of patients in need of depression care may require additional attention to improving skills of care providers. A new line of evidence indicates that a GPs emotional readiness for discussing patients' emotions might be important for the detection of depression (Baik et al., 2005). A systematic review suggests that combined with guidelines implementation, provider training offers promising results for new-onset depression patient samples (Sikorski et al., 2012).

Targeted case identification (Pignone et al., 2002, Gilbody et al., 2005, 2008) and adequate clinical skills in the recognition of patients at who may benefit from treatment (Sikorski

et al., 2012), have been argued to be of limited use if access to treatment is poor. Poor utilisation of psychological therapies in older people with arthritis and coexisting depressive disorders has previously been found in the IMPACT trial where 46% patients received any antidepressant treatment and 8% received psychological therapy in three months prior to the study (Unützer et al., 2003). A similar pattern of treatment utilisation emerged in the current study; out of 57 (39.9%) treated patients only 6 (4.2% of the whole sample, 10.5% of the treated group) patients were referred to specialist mental health professionals. When IMPACT participants were asked about preferences for depression treatment 51% reported that they would prefer counselling or psychotherapy clearly demonstrating the need for this treatment modality. Consequently, clinicians should be aware that a large proportion of patients may have a preference for psychological intervention. To address this need, GPs may use the Improving Access to Psychological Therapies programme supporting the frontline NHS scheme for implementing NICE depression and anxiety guidelines (NHS, IAPT, 2012).

As delineated in section 7.2.1 on page 228 the identification of depressive and anxiety disorders and subsequent management in chronic physical illnesses seems to pose challenges for some GPs. Given the evidence of successful implementation of collaborative depression care in OA (Lin et al., 2003, 2006) and the NICE (2009b) depression guidance for chronic physical illnesses, this model of depression care may be beneficial for patients with OA. Collaborative anxiety care in patients with OA and is still to be investigated and as yet NICE does not recommend this approach. Successful implementation of collaborative anxiety care in the CALM trial (Roy-Byrne et al., 2010), however, suggests that this model may work in patients with OA, and thus may warrant further investigation. Nevertheless,

the GP and the patient with OA may decide that the collaborative anxiety care approach would be potentially beneficial for management of anxiety symptoms.

Research implications

There are three main research implications that go beyond a need of replication of this small study in selected primary care practices. Future research would benefit from using other approaches to supplement the medical record in attempt to gauge detection. For a deeper understanding, interviews with patients and GPs may be considered. This approach may help with estimating non-documented detection, for example due to refusal to be treated or limited access to a preferred treatment modality. The role of other health professionals in the detection of depression and anxiety in older patients with OA is also still unclear. A care pathway analysis can be used in future research to gain a better understanding of their role in detection of depression and anxiety.

There is still much to do to gain a deeper understanding of why some patients with OA and depressive and/or anxiety problems are detected and not others. It seems that this issue is likely to be more complex, than easily measurable person characteristics. The effects of individual differences across primary care professionals (e.g. skills, attitude, beliefs) and patients (e.g. perception on the problem, need for care, reasons for unmet care) - on the detection of depressive or anxiety symptoms coexisting in older primary care consulters with musculoskeletal pain remain unclear. Underlying purposes of studying this problem is to gain clarity over factors associated with ‘unmet’ needs for depression and anxiety management and constituting merits of ‘adequate’ core professional skills (including diagnostic skills and personality-related factors, such as emotional readiness (Baik, 2005)).

Little is known about the consequences of non-detection. Future research may follow detected and undetected patients, to compare recovery from depressive/anxiety symptoms, mortality rates and OA-related outcomes. Research may benefit from a deeper understanding of factors associated with poor prognosis at follow-up, in particular contribution of clinical factors such as a treatment modality. This type of study may be needed to confidently recommend a need for improved identification of depressed or anxious patients.

7.8 CONCLUSIONS

Only half of all older musculoskeletal patients with persistent anxiety or depression symptoms have their mental health problems detected by their GP, despite using a definition of *detection* that includes not only Read coded consultations, but also evidence of treatment including referral and prescription. Patients frequently consulting their GPs and experiencing more severe anxiety symptoms are most likely to be detected. This study re-iterates a problem of implementing effective detection strategies for patients with painful conditions (Menchetti et al., 2009). This investigation highlights the importance of patients' awareness in their role in recognition and management of depression and anxiety problems (i.e. acknowledging the problem and self-managing it). It also highlights the need for GPs to be aware of their role in supporting patients in recognising these problems and making informed decisions about management (including the use of NHS resources). A deeper understanding of reasons underlying under-detection and effects of non-detection to patient outcomes in patients with OA still warrants further investigation.

Chapter eight: Discussion, conclusions and recommendations

8.1 INTRODUCTION

This thesis arose from a need to advance our understanding of coexisting anxiety and/or depression in patients with osteoarthritis (Katon et al., 2007, Scopaz et al. 2009). Previous studies found that depression/anxiety (RCGP, 2006) and OA are common in primary care settings (RCGP, 2007) and that they are likely to coexist (e.g. Sartorius et al., 1996, Gureje et al., 1998, Katon et al., 2007). Depressive and anxiety symptoms are known to have adverse effects on well-being of people with OA, including functioning, general health and disability (e.g. Moussavi et al., 2007, Mallen et al., 2007, Scopaz et al., 2009). According to modern psychological models of adjustment to pain (section 1.3.2 on page 12), bio-psycho-social factors are likely to be involved in the bi-directional relationship between depressive and anxiety responses to OA-related pain. Previous studies suggest that the management of depression and anxiety symptoms is challenging in older people and with chronic physical illness (Burroughs et al., 2006, Kendrick et al., 2009, Licht-Strunk et al., 2009a, Van Rijswijk et al., 2009, Coventry et al., 2011). However, effective models of depression and anxiety care exist, and can be beneficial not only for mental health care, but also for functioning, general health and disability (Lin et al., 2003, 2006, Roy-Byrne et al., 2010). The importance of the recognition and management of both depressive and anxiety symptoms coexisting with OA has been raised by many researchers and has been recently highlighted by the King's Fund and Centre for Mental Health Group (Naylor et al., 2012). This chapter presents a discussion of the work undertaken in this thesis and the contribution of this work to knowledge in this field. It starts with a summary of the main findings and is followed by a critical

reflection on the key decisions made. It ends with a consideration of the implications of this thesis for future research and clinical practice.

8.2 SUMMARY OF THE MAIN FINDINGS

The thesis started with a fundamental question about the epidemiology of depression and anxiety coexisting with OA in community-dwelling adults. For this purpose a large systematic review with meta-analyses was conducted. The findings from 36 unique original prevalence studies were synthesised, providing 'best estimates' (although based on highly heterogeneous estimates from individual studies) of frequency of a range of anxiety and depressive disorders and symptoms in people with OA. See Table 8.1 for a list of pooled 'best estimates'. These estimates of depression and anxiety concurrent with OA highlight the importance of anxiety and depression problems. Heterogeneity of these estimates reflects issues surrounding definitions, methods of ascertainment and between-study variance.

Table 8.1 Summary of pooled 'best estimates' of frequency of occurrence in people with OA.

	'Best estimate' of frequency of occurrence in people with OA (%)
Anxiety:	
Anxiety symptoms: mild or worse	45%
Anxiety symptoms: moderate or worse	21%
Generalised anxiety disorder	3%
Social phobia	3%
Panic disorder	3%
Panic with agoraphobia	2%
Post-traumatic stress disorder	2%
Depression:	
Depression symptoms: mild or worse	24%
Depression symptoms: moderate or worse	15%
Major depressive disorder	7%
Dysthymia	3%

Various self-report anxiety and depression measures are used in both UK primary care and OA research. A non-systematic review of measurement properties of several recommended patient-reported measures of depression symptoms (BDI-II and -Primary Care versions; HADS-D, PHQ-2- and -9 item versions) and anxiety (GAD-2 and -7 item versions; HADS-A) was conducted. In total, 42 articles were included. There is evidence to support some properties in some populations, but some critical properties warrant investigation, particularly in older people with OA/joint pain in the community. The feasibility and reliability of the reviewed questionnaires appears well-supported, but there is less evidence for acceptability and criterion validity. With specific reference to the HADS - a questionnaire used later in the thesis - it was concluded to perform adequately as a measure for assessing the course of depressive and anxiety symptoms (albeit the possibility of criterion contamination in the anxiety subscale). Likewise other reviewed questionnaires, point estimates based on HADS serve poor proxy for depressive and anxiety disorders, but multiple assessments can be expected to more reliably indicate anxiety and depressive problems.

Following critical consideration of the strengths and weaknesses of the secondary data source, the persistence of depressive and anxiety symptoms in a group of consecutive, older people consulting their GPs about their musculoskeletal pain was investigated. Latent class growth analysis (a person-centred method tailored for longitudinal data analysis) was used. Discrete 12-month post-consultation trajectories (also called clusters) of symptoms of anxiety and depression based on HADS anxiety and depression scores (dichotomised using a cut-off point score ≥8) were identified. These analyses were based on participants who provided complete anxiety and depression data, where no

evidence of substantial selection was found. Identified anxiety and depression symptoms trajectories are presented in Table 8.2. Persistent depression symptoms more commonly co-occurred with persistent anxiety symptoms (47%), than transient anxiety (16%) and no anxiety symptoms (7%). At baseline 48.5% and 26.5% of older patients consulting with musculoskeletal pain reported 'mild or worse' (score ≥8) anxiety and depression symptoms respectively. Persistent anxiety (30%) and depressive (22%) symptoms were common. With the aim to identify persons with persistent depressive and anxiety problems, it can be expected that 56% and 63% patients with 'mild or worse' anxiety and depression symptoms respectively at the initial presentation will have persistent anxiety and depression symptoms.

Table 8.2 Discrete trajectories of anxiety and depression symptoms nested in LCGA anxiety and depression models.

Anxiety symptoms			**Depression symptoms**		
Identified trajectories	**n**	**(%)**	**Identified trajectories**	**n**	**(%)**
No anxiety symptoms	119	(40.6%)	No depression symptoms	232	(77.8%)
Persistent anxiety symptoms	88	(30.0%)	Persistent depression symptoms	66	(22.2%)
Transient anxiety symptoms	86	(29.4%)			
Total:	293	(100%)		298	(100%)

Following careful selection of baseline variables, a group of person-characteristics were examined for associations with trajectories, using backward elimination logistic regression (for the 2-cluster LCGA depression model) and backward elimination multinomial logistic regression (for the 3-cluster LCGA

anxiety model). To improve statistical power of these analyses, sample sizes were increased. Analyses were based on samples with HADS depression scores (n=368) and anxiety scores (n=368) reported at minimum 3 time points, where each participant was assigned a cluster membership by re-estimating the selected optimal models. Pain characteristics and coping strategies were the most prominently associated with persistent depression symptoms and persistent anxiety symptoms (see Table 8.3 for a brief summary of the results of regression analyses). The identified factors have the potential to inform identification of persisting symptoms of anxiety or depression.

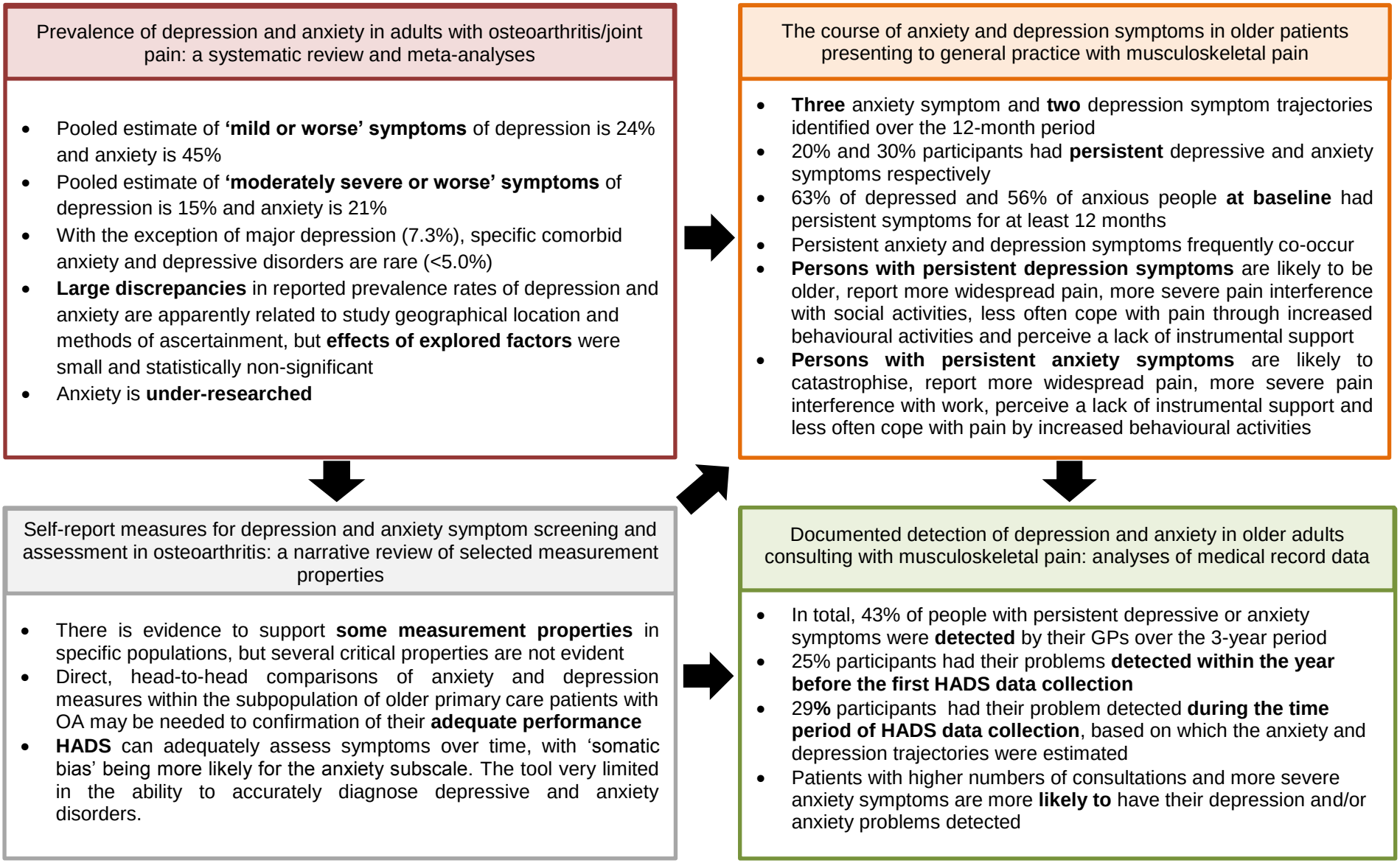
Table 8.3 Backward elimination multinomial logistic and logistic regression analyses of baseline covariates and trajectories of anxiety and depression symptoms respectively.

Baseline covariates	Anxiety trajectories		Depression trajectories
	No anxiety symptoms† vs. Persistent anxiety symptoms	**No anxiety symptoms† vs. Transient anxiety symptoms**	**No depression symptoms† vs. Persistent depression symptoms**
	Adj. OR (95%CI)	Adj. OR (95%CI)	Adj. OR (95%CI)
Age 70 years or older	-	-	3.03 (1.57, 5.84)
Catastrophising	4.14 (1.93, 8.87)	2.79 (1.30, 6.00)	1.85 (0.96, 3.55)
Coping by increased behavioural activities	0.40 (0.22, 0.75)	0.55 (0.31, 1.00)	0.51 (0.27, 0.96)
Lack of instrumental support	6.99 (1.75, 27.85)	1.54 (0.31, 7.60)	3.63 (1.32, 9.98)
Manual/routine work	1.33 (0.70, 2.54)	2.39 (1.31, 4.37)	-
Number of pain sites (0-44)	1.08 (1.0, 1.14)	1.06 (0.01, 1.11)	1.08 (1.04, 1.13)
Pain interference with social activities (0-10)	-	-	1.30 (1.14, 1.48)
Pain interference with work (0-10)	1.21 (1.08, 1.36)	1.19 (0.006, 1.33)	-

Note: † **-** a reference category.

Finally, 143 individuals identified from previous analysis as experiencing persistent anxiety and/or depression symptoms were used to investigate how often there was evidence that these problems had been detected in primary care. Data indicative of the documented *detection* of depression and anxiety was selected for extraction from the medical records, including *diagnosis/problem codes* and *interventions*. Analyses focused on a one year period following the index consultation (the main period of observation) found that 29% of patients were detected. However, as many as 25% patients had their depression and or anxiety problems detected within the year before first HADS data was collected. After the period of medical record review was extended one year either side of the main period of observation, 43% of all older musculoskeletal patients with persistent anxiety and depression symptoms had their mental health problems detected by their GP. In this sub-population of patients (as found using multivariable logistic regression analyses) those who consult more frequently (OR: 0.98, 95% CI 0.96, 0.99) and have more severe anxiety symptoms (OR: 0.82, 95%CI 0.70, 0.98) will be significantly less likely to have their depression and/or anxiety problem undetected by their GP. Figure 8.1 overleaf presents a diagram summarising the key findings of this thesis.

Figure 8.1 Diagram summarising the key finding of this thesis.



8.3 KEY DECISIONS AND THEIR IMPLICATIONS FOR THE INTERPRETATION OF THIS THESIS

In this section some of the critical overarching issues that affect the thesis and the interpretation of the findings will be discussed.

8.3.1 Definition of OA

A key issue impacting on the interpretation of the findings in thesis is the definition of osteoarthritis used. The development of a satisfactory definition of OA has presented substantial difficulties (Hurley et al., 2007). This issue also became apparent in analyses of prevalence rates conduced in chapter two. Traditional methods of defining OA, such as examination of radiographic changes or the American College of Rheumatology (Altman et al., 1986, 1990) clinical classification criteria are commonly criticised (Schouten & Valkenburg, 1995, Bierma-Zeinstra et al., 1999, Peat et al., 2006b). Evidence suggests that 'formal' diagnosis may pose a challenge to many GPs (Bierma-Zeinstra et al., 2000, Bedson et al., 2005, Peat et al., 2005). NICE (2008) uses a clinical syndrome definition, using the clinical signs and symptoms associated with OA which include use-related joint pain, joint stiffness and bony swelling associated with substantial functional limitation. In general practice, the diagnosis is typically driven by clinical judgment and is reached upon elimination of other possible causes (Altman et al., 2009). The majority of studies summarised in the systematic review included patients with joint pain, where the prevalence of depression symptoms was comparable to other definitions (with the exception of medical record based definitions). The PROG-RES participants had musculoskeletal pain that included a mixture of both OA-related and OA non-related pain reflecting the range of conditions typically encountered in primary care. Whilst many of these

patients have no ‘formal ACR’ diagnosis of OA, they are generalisable to the wider population of older people with musculoskeletal pain, many of whom will meet the clinical definition of osteoarthritis.

8.3.2 Using secondary data analyses

The use of repeated data collection over a one year period, linked to medical record data makes the PROG-RES study a good data source to answer the questions set out in this thesis. The sample itself is of limited representativeness to ethnic minority patients and for practices scoring lower than the national average Quality and Outcomes Framework score. The possibility of attrition bias was identified and the subsequent reduction in the prevalence of anxiety and depression symptoms was found. Preliminary data analyses conducted in chapter four suggested a small and statistically insignificant decrease in prevalence rates over time.

One limitation of this study is the lack of inclusion of variables that may be important in determining the relationship between OA and mental health problems. These include patient intrapersonal factors (e.g. attitudes, personality), stress factors (e.g. psychopathologies), clinical factors (e.g. effects of treatment) and qualitative aspects of socio-ecological factors on the progress of symptoms. This kind of limitation is a common problem in secondary data analyses (Boslough, 2007, Vartanian, 2010).

8.4 IMPLICATIONS

8.4.1 Implications for clinical practice

There are several direct implications of this thesis for primary care practice that will be discussed in the following section. Primary care professionals should be vigilant to the depressive/anxiety symptoms that frequently coexist in patients with OA/joint pain. Clinicians need to be prepared to face the challenge of recognising these symptoms, and formulate appropriate care plans in conjunction with their patients. Whilst the routine screening of patients with OA is currently not feasible, active case identification of those most at risk is a realistic option to identify those most in need of intervention. Focusing on persistent, rather than transient, symptoms is likely to have a greater impact on improving outcomes for those at risk of a poor outcome, although the impact of mild and moderate symptoms upon older people should not be underestimated. Healthcare professionals reported making efforts to distinguish between these two (van Rijswijk et al., 2009, Barley et al., 2011).

In UK primary care practice, as shown by Kendrick et al. (2009) self-report depression measures are likely to assist the process of identification. As shown in chapters three and five, diagnosis based on a single HADS assessment, could lead to misclassification of transient forms of anxiety and depression symptoms. Adequate measurement properties are important for assisting recognition of people who have the persistent anxiety and/or depression problem, and thus, are likely to most benefit from treatment. Chapter three was unable to confidently identify the most appropriate measure to assess anxiety and depression symptoms in older adults with OA. As the persistence of symptoms is a common characteristic, adequate responsiveness to changes in people with OA seems a

particularly important measurement property to consider (in combination with feasibility, acceptability, reliability and validity). Awareness of person-related characteristics associated with the progress of symptoms, may assist the process of identifying patients with persistent symptoms of depression and anxiety. Advantageously, primary healthcare professionals are likely to be already acknowledging the importance of a wider patient context in the diagnosis and management of the problem of depression (Dowrick et al., 2009, Barely et al. 2011). The work undertaken in this thesis supports focusing on pain-related aspects (i.e. pain extent and functional disability) and psychological factors (coping by catastrophising and increased behavioural activities).

An annual review may offer an opportunity to assess depression symptoms persistence as defined by NICE (2009b) depression guidance for chronic physical illnesses (present for a considerable time despite 'active monitoring' or low-intensity treatment). Whilst clinical guidance for anxiety in chronic physical illness is currently lacking, in seeking to recognise persistent anxiety problems GPs may choose to use the same approach as for persistent depression symptoms. 'Active monitoring', may involve the stepped approach to diagnosing depression problems in general practice, as delineated by an academic GP Lucassen et al. (2008). This theoretical model makes use of the benefit that GPs have "*over other healthcare workers because of personal continuity in a longstanding relation with the patient*" (Lucassen et al., p.164, 2008). It also emphasises the importance of hearing the patient's story, agreeing with the patient on the name of the combination of symptoms and agree on the relative importance of the problem. This view is echoed in findings of the systematic review of study of how patients understand depression with chronic physical diseases (Alderson et al., 2012). This study found a range of beliefs about

diagnostic labels, the origins of the problem and appropriateness of medication, suggesting a need for developing approaches to recognition and management, which are sensitive to a range of these beliefs (Alderson et al., 2012). As patients often view emotional representations for depression and anxiety as related (Alderson et al., 2012) they should be considered jointly, with equal attention paid to both. Healthcare professionals are likely to already have this perspective on depression and anxiety (Coventry et al., 2011, Van Rijswijk et al., 2009). Perhaps, their efforts should focus on recognising in older patients with OA those with self-perceived problems of depression and/or anxiety and a need for managing these problems. This is important as these patients are at risk of the worst course of anxiety and depression symptoms- as shown by van Beljouw et al.'s (2010) study with primary care patients. Primary care health professionals should be prepared to discuss treatment options and facilitate access to relevant health professionals. Given an apparent problem of the limited consultation time, GPs may need to collaborate with trained practice nurses or other healthcare professionals. One systematic review (Barley et al., 2011) suggested that GPs may often view practice nurses' role in recognition and management of late-life depressive disorders as limited (Burroughs et al., 2006). This approach may need to be challenged as practice nurses seem to view themselves as being in a better position to do these tasks, due to more time available and operating in a more holistic context (Murray et al., 2006). Since the recognition and treatment of late-life anxiety disorders is not supported by the QOF component of general practice contract, practice nurses' role in management of this problem may be even more limited, though this requires more investigation.

Whilst keeping in mind these important implications, the findings of this thesis should not be used to promote a predominant focus on OA-related

depression and anxiety symptoms. OA-related anxiety and depression symptoms are indicative of difficulties with adjustment to OA, and thus, should trigger co-interventions to minimise psychological and physiological symptoms and optimise functioning. There is still much to be done to improve the management of OA (Dziedzic et al., 2009). However, a good starting point may be helping patients to understand OA (including identity, impact, prognosis and current management options) and the impact of coexisting depressive/anxiety symptoms. This may be important as illness perception is known to influence coping strategies in patients in medical patients (Hagger & Orbell, 2003) and health outcomes in people with OA (Hill et al., 2007). Evidence suggests that the greatest impact of the disease lies in the effect it has on their ability to continue with a 'normal' daily life (Hill et al., 2007). To date, there is no treatment that can preserve this ability, although new promising treatments (i.e. anti-nerve growth factor (Lane et al., 2010) and strontium (Cooper et al., 2012)) have been developed. In contrast, shifting patients' perception of illness through acceptance of pain and value-based action (using acceptance-commitment therapy) (Hayes et al, 1999) was found to improve pain, depression, pain-related anxiety, disability, medical visits, work status, and physical performance (Vowles & McCracken, 2008).

One of the many challenges faced by NICE is to address a need for feasible and explicit guidance in managing psychological and physiological symptoms of OA. The challenge for policy makers is the adequate provision of required resources, including access to depression and anxiety interventions – such as acceptance-commitment therapy (Hayes et al., 1999) (offering a promising potential for reduction of both psychological and OA-related symptoms). This is essential as management of persistent depression and anxiety symptoms is unlikely to be successful without relevant interventions (Gilbody et al., 2005, 2008).

8.4.2 Implications for patients

It is perhaps true that some primary care health professionals lack knowledge and skills to recognise and manage coexisting depressive and anxiety disorders in patients with OA. Consequently their attitude reinforces patient normalising or deprioritising of depressive/anxiety symptoms. Nevertheless, it is also the role of the patient to be active in the recognition and management of their depressive and anxiety symptoms. Huber et al. (2011) recently challenged the WHO definition of health as focused on the absence of disease. The authors suggested a need for a shift towards supporting "*the ability to adapt and self-mange in the face of social, physical and emotional challenges*" (Huber et al., p.1, 2011). Although from the perspective of some GPs it is reasonable not to dismiss the WHO definition entirely (Tallini, 2011, Lewis, 2011), patient awareness of their role in management may be critical for improving the long-term prognosis.

There is ample evidence to suggest that some patients with chronic physical illnesses may find it difficult to acknowledge the need for depression help (Alderson et al., 2012). This seems to be related to a range of personal beliefs about diagnostic labels, causes, relevance and acceptability of treatment modalities social stigma, the course of symptoms or negative views about consequences of the disease (Alderson et al., 2012). Many of these patients view the GP as the right person to approach, given favourable circumstances (e.g. the patient's awareness of the problem and wish to disclose it and the GP's skills and beliefs) (Alderson et al., 2012). Evidence suggests that primary care health professionals are willing to recognise and manage the problem of depression and anxiety, despite associated difficulties (Van Rijswijk et al., 2009, Barley et al., 2011, Coventry et al., 2011). Therefore, GPs may have a unique opportunity to overcome patients' personal barriers. Nevertheless,

evidence suggests that perhaps practical issues may serve an additional obstacle for a patient to disclose the problem of depression, including limited consultation time or availability of access to preferred treatment modality [(Kadam et al., 2001)]. Since the NHS aspires to implement the principles of person-centred care, patients can assume their right to negotiate their concerns over limited consultation time and availability of access to relevant treatment. Seeking advice from a GP is one of a patient's rights, but given the organisational limitations of the NHS, it may be that a patient may need to opt for additional help from alternative sources. For example, the Arthritis Care [(Arthritis Care, 2013)] and the UK Expert Patients Programme [(Expert Patients Programme Community Interest Company, 2012)] provide and deliver free information on self-management of emotional responses to OA on a daily basis. Still, currently research cannot inform patients on acceptability and impacts of this type of support on psychological well-being and their use is likely to be depending on a patient's awareness of their existence.

8.4.3 Implications for future research

A large study with community–based adults with OA (including primary care consulters and non-consulters) is needed to compare psychometric and clinimetric properties of the HADS, BDI, PHQ and GAD. Investigated measurement properties, should include properties relevant to usefulness in identifying persistent symptoms over time (e.g. test-retest reliability, responses to changes, minimal clinically significant change and associations with OA-related outcomes). This kind of study may be designed to serve the purpose of validating discrete trajectories of anxiety and depression symptoms. Analyses should include a period of 2 years, as NICE definitions of persistent symptoms refer to this period of time, and involve

comparison of trajectories based on the HADS, PHQ, BDI and GAD measures. The most appropriate measure for primary care use should be recommended based on measurement properties and upon the evaluation of criterion validity of persistent symptoms based on this questionnaire. Analyses of factors associated with cluster membership, should include factors found to be significant in this thesis and consider some of the modifiable outcomes that were omitted in this work. These factors may be related to the GP (e.g. attitude to depression and anxiety care, emotional readiness, relevant professional skills and experience) and the patient (e.g. locus of control, attitude to mental health treatment, perceived problem and need for care). Other factors worth consideration are availability of relevant intervention, effects on OA, depression and anxiety treatments and qualitative aspects of socio-ecological factors. Next, dual anxiety and depression symptoms modeling should be conducted, including time-varying covariance with a few most prominent predictors of cluster membership or stratified by type (e.g. pain characteristics, intrapersonal factors, coping or treatment effects). Finally, this cohort may be used to evaluate contribution of factors associated with the persistence of symptoms on improved identification and impact of health outcomes.

Collaborative anxiety care warrants investigation in people with OA, using pragmatic a randomised controlled trial as opposed to an exploratory RCT. Importantly, in future studies the precise impact of issues surrounding management of depressive and anxiety symptoms need to be estimated. A study design similar to that used by van Beljouw et al. (2010) can be used to identify patients with met and un-met depression and anxiety care needs. These two groups could be then compared across potential obstacles to management (e.g.

GP-related and patient-related factors discussed in the paragraph above, availability of mental health support resources) and followed by care pathway analyses. This kind of study would help to identify the main obstacles to successful management of depression and anxiety problems in people with OA and subsequently help to optimise health care resources.

References

Abdollel M, Leblanc M, Stephensen D, Harrison RV. Binary partitioning for continuous longitudinal data: categorizing a prognostic variable. Statistics in Medicine 2002; 21(22): 3395-409.

Ahrens W, Pigeot I. Handbook of epidemiology. Berlin or Heidelberg: Springer; 2005.

Alderson SL, Foy R, Glidewell L, McLintock KL, House AO. How patients understand depression associated with chronic physical disease-a systematic review. BMC Family Practice 2012; 13(1): 41.

Ali S, Stone M, Peters J, Davies M, Khunti K. The prevalence of co-morbid depression in adults with Type 2 diabetes: a systematic review and meta-analysis. Diabetic Medicine 2006; 23(11): 1165-1173.

Allen KD, Renner JB, Devellis B, Helmick CG, Jordan JM. Osteoarthritis and sleep: the Johnston county osteoarthritis project. Journal of Rheumatology 2008; 35(6): 1102-1107.

Altman RD, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, Christy W, Cooke TD, Greenwald R, Hochberg M, Kaplan D, Koopman W, Longley S III, Mankin H, McShane DJ, Medsger T Jr, Meenan R, Mikkelsen W, Moskowitz R, Murphy W, Rothschild B, Segal M, Sokoloff L, Wolfe F. Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee. Arthritis and Rheumatism 1986; 29(8): 1039-1049.

Altman D. Practical statistics for medical research. London: Chapman and Hall; 1991.

Altman R, Alarcón G, Appelrouth D, Bloch D, Borenstein D, Brandt K, Brown C, Cooke TD, Daniel W, Feldman D, Greenwald R, Hochberg M, Howeli D, Ike R, Kapila P, Kaplan D, Koopman W, Marino C, Mcdonald E, Mcshane DJ, Medsger T, Michel b, Murphy WA, Osial T, Ramsey-Goldman R, Rothschild B, Wolfe F. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Arthritis and Rheumatism 1991; 34(5): 505-514.

Altman RD, Kuritzky L, Ruoff G. Improving long-term management of osteoarthritis: Strategies for primary care physicians. Journal of Family Practice 2009; 58(2 Suppl): S17-S24.

American Psychiatric Association (APA). Diagnostic and Statistical Manual: Mental Disorders. Washington, DC: American Psychiatric Association; 1952.

American Psychiatric Association (APA). Diagnostic and Statistical Manual of Mental Disorders, 4th ed. Washington, DC: American Psychiatric Association; 1994.

American Psychiatric Association (APA). Diagnostic and Statistical Manual of Mental Disorders, 4th ed., Text Revision. Washington, DC: American Psychiatric Association; 2000.

Anderson RJ, Freedland KE, Clouse RE, Lustman PJ. The prevalence of comorbid depression in adults with diabetes. Diabetes Care 2001; 24(6): 1069-1078.

Andrews G, Peters L. The psychometric properties of the composite international diagnostic interview. Social Psychiatry and Psychiatric Epidemiology 1998; 33(2): 80-88.

Andrews G, Slade T. The classification of anxiety disorders in ICD-10 and DSM-IV: a concordance analysis. Psychopathology 2002; 35(2-3): 100-106.

Andrews G, Slade T, Peters L. Classification in psychiatry: ICD-10 versus DSM-IV. British Journal of Psychiatry 1999; 174: 3-5.

Angst F, Verra ML, Lehmann S, Aeschlimann A. Responsiveness of five condition-specific and generic outcome assessment instruments for chronic pain. BMC Medical Research Methodology 2008; 8: 26.

Ani C, Bazargan M, Hindman D, Bell D, Farooq MA, Akhanjee L, Yemofio F, Baker R, Rodriguez M. Depression symptomatology and diagnosis: discordance between patients and physicians in primary care settings. BMC family practice 2008; 9(1): 1.

Appelt CJ, Burant CJ, Siminoff LA, Kwoh CK, Ibrahim SA. Arthritis-specific health beliefs related to aging among older male patients with knee and/or hip osteoarthritis. Journals of Gerontology Series A: Biological Sciences and Medical Sciences 2007; 62(2): 184-190.

Arbabzadeh-Bouchez S, Tylee A, Lepine JP. A European perspective on depression in the community: the DEPRES study. CNS spectrums 2002; 7: 120-128.

Arnau RC, Meagher MW, Norris MP, Bramson R. Psychometric evaluation of the Beck Depression Inventory-II with primary care medical patients. Health Psychology 2001; 20(2): 112.

Arola H, Nicholls E, Mallen C, Thomas E. Self-reported pain interference and symptoms of anxiety and depression in community-dwelling older adults: Can a temporal relationship be determined? European Journal of Pain 2010; 14(9): 966-971.

Aroll B, Khin N, Kerse N. Screening for depression in primary care with two verbally asked questions: cross-sectional study. BMJ 2003; 327: 1144-1146.

Arroll B, Goodyear-Smith F, Crengle S, Gunn J, Kerse N, Fishman T, Falloon K, Hatcher S.Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. Annals of Family Medicine 2010; 8(4): 348-353.

Arthritis Care. Understanding your feelings. 2011; Available at: http://www.arthritiscare.org.uk/LivingwithArthritis/Self-management/Understandingyourfeelings. Accessed 13/02/2013.

Arthritis and Musculoskeletal Alliance (ARMA). Standards of Care for People with Osteoarthritis. 2004; Available at: http://www.arma.uk.net/care.html. Accessed 02/11/2011.

Arthritis Research Campaign (ARC). Arthritis: the Big Picture. 2002; Available at: http://www.ipsos-mori.com/Assets/Docs/Archive/Polls/arthritis.pdf. Accessed 11/12/2010.

Asmundson GJG, Norton PJ, Vlaeyen JWS. Fear-avoidance models of chronic pain: an overview. In: Asmundson GJG, Vlaeyen JWS, Crombez G, editors. Understanding and treating fear of pain Oxford, UK: Oxford University Press; 2004. p. 3-24.

Axford J, Butt A, Heron C, Hammond J, Morgan J, Alavi A, Bolton J, Bland M. Prevalence of anxiety and depression in osteoarthritis: use of the Hospital Anxiety and Depression Scale as a screening tool. Clinical Rheumatology 2010; 29(11): 1277-1283.

Badcock L, Lewis M, Hay E, McCarney R, Croft P. Chronic shoulder pain in the community: a syndrome of disability or distress? Annals of the Rheumatic Diseases 2002; 61(2): 128-131.

Baik SY, Bowers BJ, Oakley LD, Susman JL. The recognition of depression: the primary care clinician's perspective. Annals of Family Medicine 2005; 3(1): 31-37.

Baker TA. Arthritis symptoms as indicators of pain in older African Americans. Ethnicity & Disease 2003 Fall; 13(4): 513-520.

Bakke P, Gulsvik A, Lilleng P, Overå O, Hanoa R, Eide GE. Postal survey on airborne occupational exposure and respiratory disorders in Norway: causes and consequences of non-response. Journal of Epidemiology and Community Health. 1990; 44(4): 316-320.

Barberger-Gateau P, Chaslerie A, Dartigues JF, Commenges D, Gagnon M, Salamon R. Health measures correlates in a French elderly community population: the PAQUID study. Journal of Gerontology 1992; 47(2):88-95.

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(1): 47.

Barnard KD, Skinner TC, Peveler R. The prevalence of co-morbid depression in adults with Type 1 diabetes: systematic literature review. Diabetic Med 2006; 23: 445-448.

Bauer DJ, Curran PJ. Overextraction of Latent Trajectory Classes: Much Ado About Nothing? Reply to Rindskopf (2003), Múthen (2003), and Cudeck and Henly (2003); Psychological Methods 2003; 8: 384-393.

Bay EJ, Algase DL. Fear and anxiety: a simultaneous concept analysis. International Journal of Nursing Knowledge 1999; 10(3): 103-111.

Beck AT, Steer RA. Manual for the Revised Beck Depression Inventory. San Antonio, TX: Psychological Corporation; 1993.

Beck AT, Steer RA, Ball R, Ciervo CA, Kabat M. Use of the Beck Anxiety and Depression Inventories for primary care with medical outpatients. Assessment; 4: 1997.

Beck A, Steer R, Brown G. Manual for beck depression inventory II (BDI-II). San Antonio, TX: Psychology Corporation; 1996.

Bedson J, Jordan K, Croft P. The prevalence and history of knee osteoarthritis in general practice: a case–control study. Family Practice 2005; 22(1): 103-108.

Beekman A, Deeg D, Geerlings S, Schoevers R, Smit J, Van Tilburg W. Emergence and persistence of late life depression: a 3-year follow-up of the Longitudinal Aging Study Amsterdam. Journal of Affective Disorders 2001; 65(2): 131-138.

Beesdo K, Hoyer J, Jacobi F, Low NCP, Höfler M, Wittchen HU. Association between generalized anxiety levels and pain in a community sample: evidence for diagnostic specificity. Journal of Anxiety Disorders 2009; 23(5): 684-693.

Bierma-Zeinstra S, Bohnen A, Ginai A, Prins A, Verhaar J. Validity of ACR criteria for diagnosing hip osteoarthritis in primary care research. Journal of Rheumatology 1999; 26: 1129-1133.

Bierma-Zeinstra SM, Lipschart S, Njoo KH, Bernsen R, Verhaar J, Prins A, Bohnen AM. How do general practitioners manage hip problems in adults? Scandinavian Journal of Primary Health Care 2000; 18(3): 159-164.

Blanchard RJ, Blanchard DC. Attack and defense in rodents as ethoexperimental models for the study of emotion. Progress in Neuro-Psychopharmacology & Biological Psychiatry 1989; 13: S3-S14.

Bland JM, Altman DG. Statistics notes: Cronbach's alpha. BMJ 1997; 314 (7080): 572.

Blettner M, Sauerbrei W, Schlehofer B, Scheuchenpflug T, Friedenreich C. Traditional reviews, meta-analyses and pooled analyses in epidemiology. International Journal of Epidemiology 1999; 28(1): 1-9.

BMA (British Medical Association), GPC (General Practitioners Committee). Employers National Health Service (NHS). Quality and Outcomes Framework guidance for GMS contract 2009/10. 2009; Available at: www.nhsemployers.org/Aboutus/Publications/Documents/ QOF_Guidance_2009_final.pdf. Accessed 21/07/2010.

Boardman J, Henshaw C, Willmott S. Needs for mental health treatment among general practice attenders. The British Journal of Psychiatry 2004; 185(4): 318-327.

Borenstein D. Does osteoarthritis of the lumbar spine cause chronic low back pain? Current Rheumatology Reports 2004; 6(1): 14-19.

Boslaugh S. Secondary data sources for public health: a practical guide. Cambridge: Cambridge University Press; 2007.

Bot S, Terwee C, Van der Windt D, Bouter L, Dekker J, De Vet H. Clinimetric evaluation of shoulder disability questionnaires: a systematic review of the literature. Annals of the Rheumatic Diseases 2004; 63(4): 335-341.

Bowers J, Jorm A, Henderson S, Harris P. General practitioners' detection of depression and dementia in elderly patients. Medical Journal of Australia 1990; 153(4):192.

Bowling A. research methods in health. Investigating health and health services. 2nd ed. Buckingham: Open University Press; 2002.

Brandt KD, Heilman DK, Slemenda C, Katz BP, Mazzuca S, Braunstein EM, Byrd D. A comparison of lower extremity muscle strength, obesity, and depression scores in elderly subjects with knee pain with and without radiographic evidence of knee osteoarthritis. Journal of Rheumatology 2000; 27(8): 1937.

Brosseau L, MacLeay L, Robinson V, Wells G, Tugwell P. Intensity of exercise for the treatment of osteoarthritis. Cochrane Database of Systematic Reviews 2003; (2): CD004259.

Bulat T, Castle SC, Rutledge M, Quigley P. Clinical practice algorithms: Medication management to reduce fall risk in the elderly -Part 2, summary algorithm. Journal of the American Academy of Nurse Practitioners 2008; 20(1): 1-4.

Bunevicius A, Peceliuniene J, Mickuviene N, Valius L, Bunevicius R. Screening for depression and anxiety disorders in primary care patients. Depression and Anxiety 2007; 24(7): 455-460.

Burroughs H, Lovell K, Morley M, Baldwin R, Burns A, Chew-Graham C. 'Justifiable depression': how primary care professionals and patients view late-life depression? A qualitative study. Family Practice 2006; 23(3): 369-377.

Callahan LF, Kaplan MR, Pincus T. The beck depression inventory, center for epidemiological studies depression scale (CESD), and general well-being schedule depression subscale in rheumatoid arthritis criterion contamination of responses. Arthritis and Rheumatism 1991; 4(1): 3-11.

Cameron IM, Crawford JR, Lawton K, Reid IC. Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. The British Journal of General Practice 2008; 58(546): 32.

Cameron O. The differential diagnosis of anxiety. Psychiatric and medical disorders. Psychiatric Clinics of North America 1985; 8(1): 3.

Candelas G, Abasolo L, Leon L, Lajas C, Loza E, Revenga M, Bachiller J, Collado P, Richi P, Blanco M, Jover JA. Diagnostic concordance between primary care physicians and rheumatologists in patients with work disability related to musculoskeletal disorders. Rheumatology International 2011; 31(12): 1549-1554.

Centre for Economic Performance's Mental Health Policy Group. The depression report: a new deal for depression and anxiety disorders. 2006. London: Centre for Economic Performance, The London School of Economics and Political Science.

Cesario S, Morin K, Santa-Donato A. Evaluating the Level of Evidence of Qualitative Research. Journal of Obstetric, Gynecologic, and Neonatal Nursing 2006; 31(36): 708-714.

Chatfield MD, Brayne CE, Matthews FE. A systematic literature review of attrition between waves in longitudinal studies in the elderly shows a consistent pattern of dropout between differing studies. Journal of Clinical Epidemiology 2005; 58(1): 13-19.

Ciarrochi J, Bilich L, Godsell C. Psychological flexibility as a mechanism of change in acceptance and commitment therapy. In: Ruth B, editor. Assessing Mindfulness and Acceptance: Illuminating the Processes of Change. Oakland, CA: New Harbinger Publications, Inc.; 2010. p. 51-75.

Clark DB, Jones BL, Wood DS, Cornelius JR. Substance use disorder trajectory classes: diachronic integration of onset age, severity, and course. Addictive Behaviors 2006; 31(6): 995-1009.

Clark LA, Watson D. Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications. Journal of Abnormal Psychology 1991; 100(3): 316.

Clarke M. The QUORUM statement. The Lancet 2000; 355(9205):756-757.

Cole MG. Prognosis of depression in elderly community and primary care populations: a systematic review and meta-analysis. American Journal of Psychiatry 1999; 156(8): 1182-1189.

Cole MG, Dendukuri N. Risk factors for depression among elderly community subjects: a systematic review and meta-analysis. American Journal of Psychiatry 2003; 160(6): 1147-1156.

Cook JM, Orvaschel H, Simco E, Hersen M, Joiner T. A test of the tripartite model of depression and anxiety in older adult psychiatric outpatients. Psychology and Aging 2004; 19(3): 444.

Cook AJ, Brawer PA, Vowles KE. The fear-avoidance model of chronic pain: Validation and age analysis using structural equation modeling. Pain 2006; 121(3): 195-206.

Cooper C, Reginster JY, Chapurlat R, Christiansen C, Genant H, Bellamy N, Bensen W, Navarro F, Badurski J, Nasonov E, Chevalier X, Sambrook PN. Efficacy and safety of oral strontium ranelate for the treatment of knee osteoarthritis: rationale and design of randomised, double-blind, placebo-controlled trial. Current Medical Research and Opinion 2012; 28(2): 231-239.

Corapcioglu A, Ozer GU. Adaptation of revised Brief PHQ (Brief-PHQ-r) for diagnosis of depression, panic disorder and somatoform disorder in primary healthcare settings. International Journal of Psychiatry in Clinical Practice 2004; 8(1): 11-18.

Corbière M, Bonneville-Roussy A, Franche RL, Coutu MF, Choinière M, Durand MJ, Boulanger, A. Further Validation of the BDI-II Among People With Chronic Pain Originating From Musculoskeletal Disorders. Clinical Journal of Pain 2011; 27(1): 62.

Coupland C, Dhiman P, Morriss R, Arthur A, Barton G, Hippisley-Cox J. Antidepressant use and risk of adverse outcomes in older people: population based cohort study. BMJ 2011; 343: d4551.

Coventry PA, Hays R, Dickens C, Bundy C, Garrett C, Cherrington A, Bundy C, Dickens C, Waheed W. Talking about depression: a qualitative study of barriers to managing depression in people with long term conditions in primary care. BMC family practice 2011; 12(1): 10.

Craighead WE, Nemeroff CB. The concise Corsini encyclopedia of psychology and behavioral science. 3rd ed. Hoboken, New Jersey: John Wiley and Sons; 2004.

Creamer P, Lethbridge-Cejku M, Hochberg MC. Factors associated with functional impairment in symptomatic knee osteoarthritis. Rheumatology 2000; 39(5): 490-496.

Creamer P, Lethbridge-Cejku M, Costa P, Tobin JD, Herbst JH, Hochberg MC. The relationship of anxiety and depression with self-reported knee pain in the community: Data from the Baltimore longitudinal study of aging. Arthritis Care and Research 1999; 12(1): 3-7.

Croft P, Jordan K, Jinks C. 'Pain elsewhere' and the impact of knee pain in older people. Arthritis & Rheumatism 2005; 52(8): 2350-2354.

Croghan T, Schoenbaum M, Sherbourne C, Koegel P. A framework to improve the quality of treatment for depression in primary care. Psychiatric Services 2006; 57(5): 623-630.

Croudace TJ, Jarvelin MR, Wadsworth MEJ, Jones PB. Developmental typology of trajectories to nighttime bladder control: epidemiologic application of longitudinal latent class analysis. American Journal of Epidemiology 2003; 157(9):834-842.

Cully JA, Jimenez DE, Ledoux TA, Deswal A. Recognition and treatment of depression and anxiety symptoms in heart failure. Primary care companion to the Journal of clinical psychiatry 2009; 11(3):103-109.

Dalton SO, Sorensen HT, Johansen C. SSRIs and upper gastrointestinal bleeding: what is known and how should it influence prescribing? CNS drugs 2006; 20(2): 143-151.

De Beurs E, Beekman A, Deeg D, Van Dyck R, Van Tilburg W. Predictors of change in anxiety symptoms of older persons: results from the Longitudinal Aging Study Amsterdam. Psychological Medicine 2000; 30(03): 515-527.

De Vlieger P, Crombez G, Eccleston C. Worrying about chronic pain. An examination of worry and problem solving in adults who identify as chronic pain sufferers. Pain 2006; 120(1): 138-144.

Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. BMJ 2004; 329(7458): 168-169.

Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger M, Davey Smith G, Altman DG, editors. Systematic Reviews in Health Care. Meta-analysis in Context. London: BMJ Books; 2001. p. 285-312.

Delucchi KL, Matzger H, Weisner C. Dependent and problem drinking over 5 years: a latent class growth analysis. Drug Alcohol Depend 2004; 74(3): 235-244.

Derry S, Moore RA, Rabbie R. Topical NSAIDs for chronic musculoskeletal pain in adults. Cochrane Database of Systematic Reviews 2012; (9): CD007400.

Dexter P, Brandt K. Distribution and predictors of depression symptoms in osteoarthritis. Journal of Rheumatology 1994; 21(2): 279.

Dieppe P, Brandt KD. What is important in treating osteoarthritis? Whom should we treat and how should we treat them? Rheumatic Disease Clinics of North America 2003, 29: 687–716.

Dieppe PA, Lohmander LS. Pathogenesis and management of pain in osteoarthritis. The Lancet 2005; 365(9463): 965-973.

Dionne CE, Dunn KM, Croft PR. Does back pain prevalence really decrease with increasing age? A systematic review. Age and Ageing 2006; 35(3): 229-234.

Dowrick C, Buchan I. Twelve month outcome of depression in general practice: does detection or disclosure make a difference? BMJ 1995; 311(7015): 1274-1276.

Dowrick C, Leydon GM, McBride A, Howe A, Burgess H, Clarke P, Maisey S, Kendrick T. Patients' and doctors' views on depression severity questionnaires incentivised in UK quality and outcomes framework: qualitative study. BMJ 2009; 338: b663.

Duncan R, Hay E, Saklatvala J, Croft P. Prevalence of radiographic osteoarthritis—it all depends on your point of view. Rheumatology 2006; 45(6):757-760.

Dunlop DD, Lyons JS, Manheim LM, Song J, Chang RW. Arthritis and heart disease as risk factors for major depression: the role of functional limitation. Medical Care 2004; 42(6): 502-511.

Dunlop DD, Semanik P, Song J, Manheim LM, Shih V, Chang RW. Risk factors for functional decline in older adults with arthritis. Arthritis and Rheumatism 2005; 52(4):1274-1282.

Dunn KM, Jordan K, Lacey RJ, Shapley M, Jinks C. Patterns of consent in epidemiologic research: evidence from over 25,000 responders. American Journal of Epidemiology 2004; 159(11):1087-1094.

Dunn KM, Jordan K, Croft PR. Characterizing the course of low back pain: a latent class analysis. American Journal of Epidemiology 2006; 163(8):754-761.

Dunn LB, Cooper BA, Neuhaus J, West C, Paul S, Aouizerat B, Abrams G, Edrington J, Hamolsky D, Miaskowski C. Identification of distinct depressive symptom trajectories in women following surgery for breast cancer. Health Psychology 2011; 30(6): 683–692.

Dworkin RH, Hetzel RD, Banks SM. Toward a model of the pathogenesis of chronic pain. Seminars in Clinical Neuropsychiatry 1999; 4(3): 176-185.

Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N, Carr DB, Chandler J, Cowan P, Dionne R, Galer BS, Hertz S, Jadad AR, Kramer LD, Manning DC, Martin S, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robbins W, Robinson JP, Rothman M, Royal MA, Simon L, Stauffer JW, Stein W, Tollett J, Wernicke J, Witter J. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005; 113(1-2): 9-19.

Dziedzic KS, Hill JC, Porcheret M, Croft PR. New Models for Primary Care Are Needed for Osteoarthritis. Physical Therapy 2009; 89 (12): 1371- 1378.

Eccleston C, Crombez G. Worry and chronic pain: a misdirected problem solving model. Pain 2007; 132(3): 233-236.

Egger M, Schneider M, Smith GD. Meta-analysis Spurious precision? Meta-analysis of observational studies. BMJ 1998; 316(7125): 140-144.

Egger M, Smith GD, O'Rourke K. Rationale, potentials, and promise of systematic reviews. In: Egger M, Smith GD, Altman DG, editors. Systematic reviews in health care: Meta-analysis in context. 2nd ed. London: BMJ; 2001a. p. 3-19.

Egger M, Smith GD, Sterne JAC. Uses and abuses of meta-analysis. Clinical Medicine, Journal of the Royal College of 2001b; 1(6): 478-484.

El-Rufaie O, Absood G. Retesting the validity of the Arabic version of the Hospital Anxiety and Depression (HAD) scale in primary health care. Social Psychiatry and Psychiatric Epidemiology 1995; 30(1): 26-31.

Expert Patients Programme Community Interest Company. Expert Patients Programme Courses. 2013; Available at: https://www.expertpatients.co.uk/course-participants. Accessed 13/02/2013.

Fava G, Tomba E, Sonino N. Clinimetrics: the science of clinical measurements. International Journal of Clinical Practice 2012; 66(1): 11-15.

Fazel S, Khosla V, Doll H, Geddes J. The prevalence of mental disorders among the homeless in Western countries: systematic review and meta-regression analysis. PLoS Medicine 2008; 5(12): e225.

Feldman BJ, Masyn KE, Conger RD. New approaches to studying problem behaviors: a comparison of methods for modeling longitudinal, categorical adolescent drinking data. Developmental Psychology 2009; 45(3): 652-676.

Feliciano L, Areán PA. Mood Disorders: Depressive Disorders. In: Hersen M, Turner SM, Beidel DC, editors. Adult Psychopathology and Diagnosis. 5th ed. Hoboke, New Jersey: John Wiley & Sons; 2007. p. 286-316.

Felson DT, McAlindon TE, Anderson JJ, Weissman BW, Aliabadi P, Evans S, Levy D, LaValley MP. Defining radiographic osteoarthritis for the whole knee. Osteoarthritis and Cartilage 1997; 5(4): 241-250.

Felson DT, Lawrence RC, Dieppe PA, Hirsch R, Helmick CG, Jordan JM, Kington RS, Lane NE, Nevitt MC, Zhang Y, Sowers M, McAlindon T, Spector TD, Poole AR, Yanovski SZ, Ateshian G, Sharma L, Buckwalter JA, Brandt KD, Fries JF. Osteoarthritis: new insights. Part 1: the disease and its risk factors. Annals of Internal Medicine 2000; 133(8): 635-646.

Ferreira VM, Sherman AM. Understanding associations of control beliefs, social relations, and well-being in older adults with osteoarthritis. International Journal of Aging and Human Development 2006; 62(3): 255-274.

Figaro MK, Williams-Russo P, Allegrante JP. Expectation and outlook:The impact of patient preference on arthritis care among African Americans. Journal of Ambulatory Care Management 2005; 28(1): 41-48.

First MB. Harmonisation of ICD–11 and DSM–V: opportunities and challenges. The British Journal of Psychiatry 2009; 195(5): 382-390.

Fisher MN, Al Snih S, Ostir GV, Goodwin JS. Positive affect and disability among older Mexican Americans with arthritis. Arthritis & Rheumatism-Arthritis Care & Research 2004; 51(1): 34-39.

Flink IK, Boersma K, MacDonald S, Linton SJ. Understanding catastrophizing from a misdirected problem-solving perspective. British Journal of Health Psychology 2011; 17(2): 408-419.

Fransen M, McConnell S. Exercise for osteoarthritis of the knee. Cochrane Database of Systematic Reviews 2008; (4): CD004376.

Fuller-Thomson E, Shaked Y. Factors associated with depression and suicidal ideation among individuals with arthritis or rheumatism: Findings from a representative community survey. Arthritis Care and Research 2009; 61(7): 944-950.

Furukawa TA, Guyatt GH. Sources of bias in diagnostic accuracy studies and the diagnostic process. Canadian Medical Association Journal 2006; 174(4): 481-482.

García-Campayo J, Zamorano E, Ruiz MA, Pardo A, Pérez-Páramo M, López-Gómez V, Freire O, Rejas J. Cultural adaptation into Spanish of the generalized anxiety disorder-7 (GAD-7) scale as a screening tool. Health and Quality of Life Outcomes 2010; 8(8): 1-11.

Gerrits MMJG, Vogelzangs N, van Oppen P, van Marwijk HWJ, van der Horst H, Penninx BWJH. Impact of pain on the course of depressive and anxiety disorders. Pain 2012; 153(2): 429-36.

Gibson C. Semi-structured and unstructured interviewing: a comparison of methodologies in research with patients following discharge from an acute psychiatric hospital. Journal of Psychiatric and Mental Health Nursing 1998; 5(6): 469-477.

Gignac M, Backman CL, Davis AM, Lacaille D, Mattison CA, Montie P, Badley EM. Understanding social role participation: what matters to people with arthritis? Journal of Rheumatology 2008; 35(8): 1655-1663.

Gilbody S, House AO, Sheldon TA. Routinely administered questionnaires for depression and anxiety: systematic review. BMJ 2001; 322(7283): 406-409.

Gilbody S, House A, Sheldon T. Screening and case finding instruments for depression. Cochrane Database Systematic Review 2005; 4: CD002792.

Gilbody S, Sheldon T, House A. Screening and case-finding instruments for depression: a meta-analysis. Canadian Medical Association Journal 2008; 178(8): 997-1003.

Gilchrist G, Gunn J. Observational studies of depression in primary care: what do we know? BMC Family practice 2007; 8(1): 28.

Glass GE. Update: spatial aspects of epidemiology: the interface with medical geography. Epidemiologic Reviews 2000; 22(1): 136-139.

Glasziou P. Systematic reviews in health care: a practical guide. Cambridge: Cambridge University Press; 2001.

GL Education Group. Hospital Anxiety and Depression Scale. 2012; Available at: http://www.gl-assessment.co.uk/products/hospital-anxiety-and-depression-scale-0 Accessed 20/01/2012.

Glover G, Webb M, Evison F. Improving Access to Psychological Therapies A review of the progress made by sites in the first roll-out year. The North East Public Health Observatory 2010.

Goldberg DP, Lecrubier Y. Form and frequency of mental disorders across centres. In: Form TB, Sartorius N, editors. Mental Illness in General Health Care. An International Study. Wiley Publishers: Chichester; 1995. p. 323-334.

Gould W. Sg124: interpreting logistic regression in all its forms. STATA Technical Bulletin 2000; 53: 18-29.

Goodman JH, Tyer-Viola L. Detection, treatment, and referral of perinatal depression and anxiety by obstetrical providers. Journal of Women's Health 2010; 19(3): 477-490.

Griffin JM, Fuhrer R, Stansfel SA, Marmot M. Theimportance of low control at work and home on depression and anxiety: Do these effects vary by gender and social class? Social Science & Medicine 2002; 54: 783–798.

Grigsby AB, Anderson RJ, Freedland KE, Clouse RE, Lustman PJ. Prevalence of anxiety in adults with diabetes: A systematic review. Journal of Psychosomatic Research 2002 12; 53(6): 1053-1060.

Grimes DA, Schulz KF. Uses and abuses of screening tests. The Lancet 2002; 359(9309): 881-884.

Gruenberg AM, Goldstein RD, Pincus HA. Classification of Depression: Research and Diagnostic Criteria: DSM-IV and ICD-10. Biology of Depression 2005:1-12. In: Licinio J, Wong ML, editors. Biology of depression. From novel insights to therapeutic strategies. Weinheim, Germany: Wiley-VCH Verlag GmbH and Company; 2005. p. 1-12.

Gureje O, Von Korff M, Simon GE, Gater R. Persistent pain and well-being. JAMA: the journal of the American Medical Association 1998; 280(2): 147-151.

Gureje O, Ademola A, Olley BO. Depression and disability: comparisons with common physical conditions in the Ibadan Study of Aging. Journal of the American Geriatrics Society 2008; 56(11): 2033-2038.

Hagger MS, Orbell S. A meta-analytic review of the common-sense model of illness representations. Psychology and Health 2003; 18(2): 141-184.

Hampson SE, Glasgow RE, Zeiss AM. Coping with osteoarthritis by older adults. Arthritis and Rheumatism 1996; 9(2):133-141.

Han C, Jo SA, Kwak JH, Pae CU, Steffens D, Jo I, Park MH. Validation of the Patient Health Questionnaire-9 Korean version in the elderly population: the Ansan Geriatric study. Comprehensive Psychiatry 2008; 49(2): 218-223.

Hanel G, Henningsen P, Herzog W, Sauer N, Schaefert R, Szecsenyi J, Löwe B. Depression, anxiety, and somatoform disorders: Vague or distinct categories in primary care? Results from a large cross-sectional study. Journal of Psychosomatic Research 2009 9; 67(3): 189-197.

Hannan MT, Felson DT, Pincus T. Analysis of the discordance between radiographic changes and knee pain in osteoarthritis of the knee. Journal of Rheumatology 2000; 27(6): 1513-1517.

Harbord RM, Higgins JPT. Meta-regression in Stata. Stata Journal 2008; 8(4): 493-529.

Harbord RM, Harris RJ, Sterne JAC. Updated tests for small-study effects in meta-analyses. Stata Journal 2009; 9(2): 197-210.

Harris T, Cook DG, Victor C, DeWilde S, Beighton C. Onset and persistence of depression in older people - results from a 2-year community follow-up study. Age and Ageing 2006; 35(1): 25-32.

Harris CA, D'Eon JL. Psychometric properties of the Beck Depression Inventory-Second Edition (BDI-II) in individuals with chronic pain. Pain 2008; 137(3): 609-622.

Hawker G, Stewart L, French M, Cibere J, Jordan J, March L, Suarez-Almazor M. Understanding the pain experience in hip and knee osteoarthritis - an OARSI/OMERACT initiative. Osteoarthritis and Cartilage 2008; 16(4): 415-422.

Hayes SC, Strosahl KD, Wilson KG. Acceptance and commitment therapy: An experiential approach to behavior change. New York: Guilford Press; 1999.

Haythornthwaite JA, Menefee LA, Heinberg LJ, Clark MR. Pain coping strategies predict perceived control over pain. Pain 1998; 77(1): 33-39.

Hays R, Hadorn D. Responsiveness to change: an aspect of validity, not a separate dimension. Quality of Life Research 1992; 1(1): 73-75.

He Y. Missing data analysis using multiple imputation. Circulation: Cardiovascular Quality and Outcomes 2010; 3(1): 98-105.

He Y, Zhang M, Lin EHB, Bruffaerts R, Posada-Villa J, Angermeyer M, Levinson D, de Girolamo G, Uda H, Mneimneh Z, Benjet C, de Graaf R, Scott KM, Gureje O, Seedat S, Haro JM, Bromet EJ, Alonso J, von Korff M, Kessler RC. Mental disorders among persons with arthritis: results from the World Mental Health Surveys. Psychological Medicine 2008; 38(11): 1639.

Henkel V, Mergl R, Coyne JC, Kohnen R, Möller HJ, Hegerl U. Screening for depression in primary care: Will one or two items suffice? European Archives of Psychiatry and Clinical Neuroscience 2004; 254(4): 215-223.

Heuts PHTG, Vlaeyen JWS, Roelofs J, de Bie RA, Aretz K, van Weel C, van Schayck OC. Pain-related fear and daily functioning in patients with osteoarthritis. Pain 2004; 110(1-2): 228-235.

Hickie IB, Andrews G, Davenport TA. Measuring outcomes in patients with depression or anxiety: an essential part of clinical practice. Medical Journal of Australia 2002; 177(4): 205-207.

Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003; 327(7414): 557-560.

Higgins JPT. Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. International Journal of Epidemiology 2008; 37(5): 1158-1160.

Hill S, Dziedzic K, Thomas E, Baker SR, Croft P. The illness perceptions associated with health and behavioural outcomes in people with musculoskeletal hand problems: Findings from the North Staffordshire Osteoarthritis Project (NorStOP). Rheumatology 2007; 46(6): 944-951.

Hipp JR, Bauer DJ. Local solutions in the estimation of growth mixture models. Psychological Methods 2006; 11(1): 36.

Hippisley-Cox J, Jumbu G. Trends in Consultation Rates in General Practice 1995/1996 to 2007/2008: analysis of the QRESEARCH database; 2008.

Hopman-Rock M, Odding E, Hofman A, Kraaimaat FW, Bijlsma J. Differences in health status of older adults with pain in the hip or knee only and with additional mobility restricting conditions. Journal of Rheumatology 1997; 24(12): 2416.

Huber M, Knottnerus JA, Green L, van der Horst H, Jadad AR, Kromhout D, Leonard B, Lorig K, Loureiro MI, van der Meer JW, Schnabel P, Smith R, van Weel C, Smid H. How should we define health? British Medical Journal 2011: 343: d4163.

Hunsley J, Mash EJ. The role of assessment in evidence-based practice. In: Antony MM, Barlow DH, editors. Handbook of assessment and treatment planning for psychological disorders. 2nd ed. New York: The Guilford Press; 2010. p. 3-22.

Hurley M, Walsh N, Mitchell H, Pimm T, Patel A, Williamson E, Jones R, Reeves B, Dieppe P. Clinical effectiveness of a rehabilitation program integrating exercise, self-management, and active coping strategies for chronic knee pain: A cluster randomized trial. Arthritis Care and Research 2007; 57(7): 1211-1219.

Hutton C, Vennart W. Osteoarthritis and magnetic resonance imaging: potential and problems. Annals of the Rheumatic 1995; 54(4): 237-243.

Iliffe S, Haines A, Gallivan S, Booroff A, Goldenberg E, Morgan P. Assessment of elderly people in general practice. 1. Social circumstances and mental state. British Journal of General Practice 1991; 41(342): 9-12.

Iliffe S, Kharicha K, Carmaciu C, Harari D, Swift C, Gillman G, Stuck AE. The relationship between pain intensity and severity and depression in older people: exploratory study. BMC family practice 2009; 10(1): 54.

Jacobsen PB, Donovan KA, Swaine ZN, Watson IS. Management of anxiety and depression in adult cancer patients: Toward an evidence-based approach. In: Chang AE, Ganz PA, Hayes DF, Kinsella T, Pass HI, Schiller JH, Stone R, Strecher V. editors. Oncology: An evidence-based approach. New York, NY: Springer-Verlag; 2006. p. 1561-1588.

Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA 1994; 271(9):703–707

Jakobsson U, Hallberg IR. Quality of life among older adults with osteoarthritis: an explorative study. Journal of Gerontological Nursing 2006; 32(8): 51-60.

Jensen MP, Keefe FJ, Lefebvre JC, Romano JM, Turner JA. One-and two-item measures of pain beliefs and coping strategies. Pain 2003; 104(3): 453-469.

Jeste DV, Blazer DG, First M. Aging-related diagnostic variations: need for diagnostic criteria appropriate for elderly psychiatric patients. Biological Psychiatry 2005; 58(4): 265-271.

Joint Formulary Committee. British National Formulary (BNF). 6th ed. London: BMJ Group and Pharmaceutical Press; 2011.

Joling KJ, van Marwijk HWJ, Piek E, der Horst HE, Penninx BW, Verhaak P, van Hout HPJ. Do GPs' medical records demonstrate a good recognition of depression? A new perspective on case extraction. Journal of Affective Disorders 2011; 133(3): 522-527.

Jones CM, Athanasiou T. Summary receiver operating characteristic curve analysis techniques in the evaluation of diagnostic tests. Annals of Thoracic Surgery 2005; 79(1): 16-20.

Jooste P, Yach D, Steenkamp H, Botha J, Rossouw J. Drop-out and newcomer bias in a community cardiovascular follow-up study. International journal of epidemiology 1990; 19(2): 284-289.

Jordan KP, Thomas E, Peat G, Wilkie R, Croft P. Social risks for disabling pain in older people: a prospective study of individual and area characteristics. Pain 2008; 137(3): 652-661.

Julian LJ. Measures of anxiety: State-Trait Anxiety Inventory (STAI), Beck Anxiety Inventory (BAI), and Hospital Anxiety and Depression Scale-Anxiety (HADS-A). Arthritis Care and Research 2011; 63(S11): S467-S472.

Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA: the journal of the American Medical Association 1999; 282(11): 1054-1060.

Kadam UT, Croft P, McLeod J, Hutchinson M. A qualitative study of patients' views on anxiety and depression. British Journal of General Practice 2001; 51(466): 375.

Kadam UT, Croft PR. Clinical comorbidity in osteoarthritis: Associations with physical function in older patients in family practice. Journal of Rheumatology 2007; 34(9): 1899-1904.

Kalichman L, Li L, Kim D, Guermazi A, Berkin V, O'Donnell CJ, Hoffmann U, Cole R, Hunter DJ. Facet joint osteoarthritis and low back pain in the community-based population. Spine 2008; 33(23): 2560.

Kalichman L, Zhu Y, Zhang Y, Niu J, Gale D, Felson DT, Hunter DJ. The association between patella alignment and knee pain and function: an MRI study in persons with symptomatic knee osteoarthritis. Osteoarthritis and Cartilage 2007; 15(11): 1235-1240.

Kalin NH. The neurobiology of fear. Scientific American 1993; 268(5): 94-101.

Kaptein KI, de Jonge P, van den Brink RHS, Korf J. Course of depression symptoms after myocardial infarction and cardiac prognosis: a latent class analysis. Psychosomatic Medicine 2006; 68(5): 662-668.

Karp DA. Living with depression: illness and identity turning points. Qualitative Health Research 1994; 4 (1): 6-30.

Kass RE, Raftery AE. Bayes factors and model uncertainty. Journal of the American Statistical Association 1995; 90(430): 773-795.

Katon W, Lin EHB, Kroenke K. The association of depression and anxiety with medical symptom burden in patients with chronic medical illness. General Hospital Psychiatry 2007; 29(2): 147-155.

Kendell RC. The role of diagnosis in psychiatry. Oxford, England: Blackwell Scientific; 1975.

Kendrick T. Depression in adults: GPs are not so bad at diagnosis. BMJ 2008; 336(7643): 522.

Kendrick T, Dowrick C, McBride A, Howe A, Clarke P, Maisey S, Moore M, Smith PW. Management of depression in UK general practice in relation to scores on depression severity questionnaires: analysis of medical record data. BMJ 2009; 338: b750.

Kendrick T, Peveler R. Guidelines for the management of depression: NICE work? The British Journal of Psychiatry 2010; 197(5):345-347.

Kessler D, Bennewith O, Lewis G, Sharp D. Detection of depression and anxiety in primary care: follow up study. BMJ 2002; 325(7371): 1016-1017.

Kessler D, Sharp D, Lewis G. Screening for depression in primary care. The British Journal of General Practice 2005; 55(518): 659-660.

Khan KS, ter Riet G, Glanville J, Sowden AJ, Kleijnen J. Undertaking systematic reviews of research on effectiveness: CRD's guidance for carrying out or commissioning reviews. 2nd ed. CRD Report No. 4. York: NHS Centre for Reviews and Dissemination; 2001.

Kim KW, Han JW, Cho HJ, Chang CB, Park JH, Lee JJ, Lee SB, Seong SC, Kim TK. Association between comorbid depression and osteoarthritis symptom severity in patients with knee osteoarthritis. The Journal of Bone and Joint Surgery (American) 2011; 93(6): 556-563.

Klapow J, Kroenke K, Horton T, Schmidt S, Spitzer R, Williams JBW. Psychological disorders and distress in older primary care patients: a comparison of older and younger samples. Psychosomatic Medicine 2002; 64(4): 635-643.

Klassen TP, Jadad AR, Moher D. Guides for reading and interpreting systematic reviews: I. Getting started. Archives of Pediatrics and Adolescent Medicine 1998; 152(7): 700.

Kline P. An easy guide to factor analysis. London: Routledge; 1994.

Kopec JA, Rahman MM, Sayre EC, Cibere J, Flanagan WM, Aghajanian J, Anis AH, Jordan JM, Badley EM. Trends in physician-diagnosed osteoarthritis incidence in an administrative database in British Columbia, Canada, 1996–1997 through 2003–2004. Arthritis Care and Research 2008; 59(7): 929-934.

Kovacs F, Noguera J, Abraira V, Royuela A, Cano A, Gil del Real MT, Zamora J, Gestoso M, Muriel A, Mufraggi N. The Influence of Psychological Factors on Low Back Pain-Related Disability in Community Dwelling Older Persons. Pain Medicine 2008; 9(7): 871-880.

Kramer SE, Kapteyn TS, Kuik DJ, Deeg DJH. Association of hearing impairment and chronic diseases with psychosocial health status in older age. Journal of Aging and Health 2002; 14(1): 122-137.

Kratz AL. Pain acceptance moderates the relation between pain and negative affect in female osteoarthritis and fibromyalgia patients. Annual Behavioural Medicine 2007; 33: 291.

Kreuter F, Muthén B. Analyzing Criminal Trajectory Profiles: Bridging Multilevel and Group-based Approaches Using Growth Mixture Modeling Journal of Quantitative Criminology 2008; 24(1):1-31.

Kroenke K, Spitzer RL, Williams JBW. The PHQ-9. A validation of a brief depression severity measure. Journal of General Internal Medicine 2001; 16(9): 606-613.

Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: validity of a two-item depression screener. Medical Care 2003; 41(11): 1284.

Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Annals of Internal Medicine 2007; 146(5): 317-325.

Krugh HE, Woods SR, Mahowald ML. The importance of identifying depression in patients with rheumatoid arthritis: evaluation of the beck depression inventory. JCR: Journal of Clinical Rheumatology1997; 3(5): 248.

Krumholz HM, Butler J, Miller J, Vaccarino V, Williams CS, de Leon CFM, Seeman TE, Kasl SV, Berkman LF. Prognostic importance of emotional support for elderly patients hospitalized with heart failure. Circulation 1998; 97(10): 958-964.

Kumar R, Indrayan A. Receiver operating characteristic (ROC) curve for medical researchers. Indian Pediatrics 2011; 48(4): 277-287.

Lacey RJ, Lewis M, Jordan K, Jinks C, Sim J. Interrater reliability of scoring of pain drawings in a self-report health survey. Spine 2005; 30(16): E455-8.

Lam CLK, Pan PC, Chan AWT, Chan SY, Munro C. Can the Hospital Anxiety and Depression (HAD) Scale be used on Chinese elderly in general practice? Family Practice 1995; 12(2): 149-154.

Lane NE, Schnitzer TJ, Birbara CA, Mokhtarani M, Shelton DL, Smith MD, Brown MT. Tanezumab for the treatment of pain from osteoarthritis of the knee. New England Journal of Medicine 2010; 363(16): 1521-1531.

Lane NE, Brandt K, Hawker G, Peeva E, Schreyer E, Tsuji W, Hochberg MC. OARSI-FDA initiative: defining the disease state of osteoarthritis. Osteoarthritis and Cartilage / OARS, Osteoarthritis Research Society 2011; 19(5): 478-482.

Lang AJ, Stein MB. Screening for anxiety in primary care: why bother? General Hospital Psychiatry 2002; 24(6): 365-366.

Laursen B, Hoff E. Person-Centered and Variable-Centered Approaches to Longitudinal Data. Merrill-Palmer Quarterly: Journal of Developmental Psychology 2006; 52: 377-389.

Lazarus RS. Stress and emotion: A new synthesis. Springer Publishing Company; 2006.

LeDoux J. The emotional brain: the mistirious underpinnings of emotional life. New York: Simon & Schuster; 1996.

LeDoux J. The emotional brain, fear, and the amygdala. Cellular and molecular neurobiology 2003; 23(4):727-738.

Lee GK, Chan F, Berven NL. Factors affecting depression among people with chronic musculoskeletal pain: A structural equation model. Rehabilitation Psychology 2007; 52(1): 33-43.

Lee GK, Mercurio-Riley D. Psychosocial Adjustment of People with Chronic Pain: A Conceptual Framework of Risk and Resistance Factors. Australian Journal of Rehabilitation Counselling 2009; 15(1): 40-60.

Leeflang MMG, Deeks JJ, Gatsonis C, Bossuyt PMM. Systematic reviews of diagnostic test accuracy. Annals of Internal Medicine 2008; 149(12): 889-897.

Lester H, Glasby J. Mental health policy and practice. Basingstoke: Palgrave Macmillan; 2006.

Leveille SG, Bean J, Ngo L, McMullen W, Guralnik JM. The pathway from musculoskeletal pain to mobility difficulty in older disabled women. Pain 2007; 128(1-2): 69-77.

Levenson JLE. Psychiatric care for medically ill. In: Levenson JL, editor. The American psychiatric publishing textbook of psychometric medicine. 2nd ed. Washington C.C.: The American Psychiatric Publishing, Inc.; 2010. p. 241.

Lewinnek GE, Warfield CA. Facet joint degeneration as a cause of low back pain. Clinical Orthopedy 1986; 213: 216–222.

Lewis M, Jordan K, Lacey R, Jinks C, Sim J. Inter-rater reliability assessment of the scoring of the body pain manikin. Rheumatology 2002; 41: 52.

Lewis DM. How should we define health? WHO definition of health remains fit for purpose. BMJ 2011; 343:d5357

Li C, Friedman B, Conwell Y, Fiscella K. Validity of the Patient Health Questionnaire 2 (PHQ-2) in Identifying Major Depression in Older People. Journal of the American Geriatrics Society 2007; 55(4): 596-602.

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis J, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Journal of Clinical Epidemiology 2009; 62(10): e1-e34.

Licht-Strunk E, Van Der Windt DAWM, Van Marwijk HWJ, De Haan M, Beekman ATF. The prognosis of depression in older patients in general practice and the community. A systematic review. Family Practice 2007; 24(2): 168-180.

Licht-Strunk E. The Prognosis of Late-Life Depression in General Practice. Amsterdam: VU University Medical Centre Amsterdam; 2008.

Licht-Strunk E, Beekman ATF, de Haan M, van Marwijk HWJ. The prognosis of undetected depression in older general practice patients. A one year follow-up study. Journal of Affective Disorders 2009a; 114(1-3): 310-315.

Licht-Strunk E, Van Marwijk H, Hoekstra T, Twisk J, De Haan M, Beekman A. Outcome of depression in later life in primary care: longitudinal cohort study with three years' follow-up. BMJ 2009b; 338: a3079.

Lin E, Katon W, Von Korff M, Tang L, Williams Jr JW, Kroenke K, Hunkeler E, Harpole L, Hegel M, Arean P, Hoffing M, Della Penna R, Langston C, Unützer J. Effect of improving depression care on pain and functional outcomes among older adults with arthritis. JAMA 2003; 290(18): 2428-2429.

Lin E, Tang L, Katon W, Hegel MT, Sullivan MD, Unützer J. Arthritis pain and disability: response to collaborative depression care. General Hospital Psychiatry 2006; 28(6): 482-486.

Lipsey MW, Wilson DB. Practical meta-analysis. Thousand Oaks, CA: Sage Publications, Inc; 2001.

Loney PL, Stratford PW. The Prevalence of Low Back Pain in Adults: A Methodological Review of the Literature. Physical Therapy 1999; 79(4): 384-396.

López-Ibor J, Frances A, Jones C. Dysthymic disorder: a comparison of DSM-IV and ICD-10 and issues in differential diagnosis. Acta Psychiatrica Scandinavica 1994; 89: 12-18.

López-Ibor JJ. The classification of stress-related disorders in ICD-10 and DSM-IV. Psychopathology 2002; 35(2-3): 107-111.

Lopez-Lopez A, Montorio I, Izal M, Velasco L. The role of psychological variables in explaining depression in older people with chronic pain. Aging and Mental Health 2008; 12(6): 735-745.

Löwe B, Kroenke K, Herzog W, Gräfe K. Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). Journal of Affective Disorders 2004a; 81(1): 61-66.

Löwe B, Unützer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the patient health questionnaire-9. Medical Care 2004b; 42(12): 1194.

Lucassen P, Van Rijswijk E, van Weel-Baumgarten E, Dowrick C. Making fewer depression diagnoses: beneficial for patients? Mental Health in Family Medicine 2008; 5(3): 161.

Luppa M, Sikorski C, Luck T, Ehreke L, Konnopka A, Wiese B, Weyerer S, König HH, Riedel-Heller SG. Age-and gender-specific prevalence of depression in latest-life-Systematic review and meta-analysis. Journal of Affective Disorders 2010; doi:10.1016/j.jad.2010.11.033.

MacMillan HL, Patterson CJS, Wathen CN. Screening for depression in primary care: recommendation statement from the Canadian Task Force on Preventive Health Care. Canadian Medical Association Journal 2005; 172(1): 33-35.

Magidson J, Vermunt J. Latent class models. In: Kaplan D, editor. The SAGE handbook of quantitative methodology for the social sciences. Thousand Oakes: Sage Publications; 2004. p. 175-198.

Mallen C, Peat G, Croft P. Quality assessment of observational studies is not commonplace in systematic reviews. Journal of Clinical Epidemiology 2006a 8; 59(8): 765-769.

Mallen CD, Peat G, Thomas E, Wathall S, Whitehurst T, Clements C, Bailey J, Gray J, Croft PR. The assessment of the prognosis of musculoskeletal conditions in older adults presenting to general practice: a research protocol. BMC musculoskeletal disorders 2006b; 7(1): 84.

Mallen CD, Peat G, Thomas E, Lacey R, Croft P. Predicting poor functional outcome in community-dwelling older adults with knee pain: prognostic value of generic indicators. Annals of Rheumatic Disease 2007; 66(11): 1456-1461.

Mallen CD, Peat G. Screening older people with musculoskeletal pain for depression symptoms in primary care. British Journal Of General Practice 2008 10; 58(555):688-693.

Mallen CD. The progress of musculoskeletal pain in older adults in general practice. Keele University: Arthritis Research UK Primary Care Centre; 2009.

Maly MR, Costigan PA, Olney SJ. Self-efficacy mediates walking performance in older adults with knee osteoarthritis. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences 2007; 62(10): 1142-1146.

Marengoni A, Angleman S, Melis R, Mangialasche F, Karp A, Garmen A, Meinow B, Fratiglioni L. Aging with multimorbidity: A systematic review of the literature. Ageing Research Reviews 2011; 10(4): 430-439.

Marshall GN, Sherbourne CD, Meredith LS, Camp P, Hays RD. The tripartite model of anxiety and depression: symptom structure in depressive and hypertensive patient groups. Journal of Personality Assessment 2003; 80(2): 139-153.

Martin JC. Determinants of functional health of low-income black women with osteoarthritis. American Journal of Preventive Medicine 1996; 12(5): 430-436.

McAlindon T, Zhang Y, Hannan M, Naimark A, Weissman B, Castelli W, Felson D. Are risk factors for patellofemoral and tibiofemoral knee osteoarthritis different? Journal of rheumatology 1996; 23(2): 332.

McCormick A, Fleming D, Charlton J. Morbidity Statistics from General Practice. Fourth National Study 1991–1992. Office of Population Censuses and Surveys. 1995; Series MB5 No.3. McCracken LM, Eccleston C. Coping or acceptance: what to do about chronic pain? Pain 1995; 105(1-2):197-204.

McCracken LM, Spertus IL, Janeck AS, Sinclair D, Wetzel FT. Behavioral dimensions of adjustment in persons with chronic pain: pain-related anxiety and acceptance. Pain 1999; 80(1):283-289.

McLachlan GJ, Peel D. Finite mixture models. New York: Wiley-Interscience; 2000.

McWilliams LA, Goodwin RD, Cox BJ. Depression and anxiety associated with three pain conditions: results from a nationally representative sample. Pain 2004; 111(1-2):77-83.

McWilliams LA, Clara IP, Murphy PDJ, Cox BJ, Sareen J. Associations between arthritis and a broad range of psychiatric disorders: Findings from a nationally representative sample. Journal of Pain 2008; 9(1): 37-44.

Meades R, Ayers S. Anxiety measures validated in perinatal populations: A systematic review. Journal of Affective Disorders 2010; 133(1-2): 1-15.

Memel DS, Kirwan JR, Sharp DJ, Hehir M. General practitioners miss disability and anxiety as well as depression in their patients with osteoarthritis. The British Journal Of General Practice 2000 08; 50(457): 645-648.

Menchetti M, Murri MB, Bertakis K, Bortolotti B, Berardi D. Recognition and treatment of depression in primary care: effect of patients' presentation and frequency of consultation. Journal of Psychosomatic Research 2009; 66(4): 335-341.

Mendels, J. Concepts of Depression. New York: Wiley; 1970.

Menz H, Tiedemann A, Kwan M, Plumb K, Lord SR. Foot pain in community-dwelling older people: an evaluation of the Manchester Foot Pain and Disability Index. Rheumatology 2006; 45(7): 863-867.

Miller LR, Cano A. Comorbid chronic pain and depression: who is at risk? The Journal of Pain 2009; 10(6): 619-627.

Mitchell AJ, Coyne JC. Do ultra-short screening instruments accurately detect depression in primary care?: A pooled analysis and meta-analysis of 22 studies. British Journal of General Practice 2007; 57(535): 144-151.

Mitchell AJ. A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. Journal of Psychiatric Research 2009; 43(4): 411-431.

Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. The Lancet 2009; 374(9690): 609-619.

Mitchell AJ, Chan M, Bhatti H, Halton M, Grassi L, Johansen C, Meader N. Prevalence of depression, anxiety, and adjustment disorder in oncological, haematological, and palliative-care settings: a meta-analysis of 94 interview-based studies. The Lancet Oncology 2011 2; 12(2):160-174.

Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. Journal of Clinical Epidemiology 2010; 63(7): 737-745.

Montgomery S. Anxiety and Depression. London: Livingstone; 1990.

Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. The Lancet 2007; 370(9590): 851-858.

Munce SEP, Stewart DE. Gender differences in depression and chronic pain conditions in a national epidemiologic survey. Psychosomatics 2007; 48(5): 394-399.

Murray J, Banerjee S, Byng R, Tylee A, Bhugra D, Macdonald A. Primary care professionals' perceptions of depression in older people: a qualitative study. Social Science and Medicine 2006; 63(5):1363-1373.

Muthén L, Muthén B. Mplus 5.21: Statistical analysis with latent variables [computer software]. Los Angeles, CA: Muthén & Muthén; 2009.

Muthén B, Muthén LK. Integrating Person-Centered and Variable-Centered Analyses: Growth Mixture Modeling With Latent Trajectory Classes. Alcoholism: Clinical and Experimental Research 2000; 24(6): 882-891.

Muus KJ, McDonald LR, Ludtke RL, Allery AJ, Knudson A, Moulton P. Arthritis among American Indian and Alaska Native elders: prevalence, demographic patterns, and comorbidities. Journal of Native Aging and Health 2007; 2(1): 5-13.

Myers K, Winters N. Ten-year review of rating scales. I: Overview of scale functioning, psychometric properties, and selection. Journal of American Academy of Childhood and Adolescent Psychiatry 2002; 41 (2):114-121.

Nagin D. Group-based modeling of development. Cambridge, MA: Harvard University Press; 2005.

Nagin DS, Tremblay RE. Analyzing developmental trajectories of distinct but related behaviors: A group-based method. Psychological Methods 2001; 6(1): 18-34.

Nandi A, Beard J, Galea S. Epidemiologic heterogeneity of common mood and anxiety disorders over the life course in the general population: a systematic review. BMC Psychiatry 2009; 9(1): 31.

NCC-CC (National Collaborating Centre for Chronic Conditions). Osteoarthritis: national clinical guideline for care and management in adults. 2008; 90.

NCC-MH (National Collaborating Centre for Mental Health). Post-traumatic stress disorder. The management of PTSD in adults and children in primary and secondary care. 2005; 26.

NHS (National Health Service), IAPT (Improving Access to Psychological Therapies). Relieving distress, transforming lives. 2013; Available at: http://www.iapt.nhs.uk/. Accessed 13/12/2012.

NHSIA (NHS Information Authority). Towards an Information Standard for Organising Clinical Communications. London: NHS Information Authority; 2000.

NICE (National Institute for Health and Clinical Excellence). Depression: management of depression in primary and secondary care. 2004; 23.

NICE (National Institute for Health and Clinical Excellence). Obsessive-compulsive disorder: core interventions in the treatment of obsessive-compulsive disorder and body dysmorphic disorder. 2005; 31.

NICE (National Institute for Health and Clinical Excellence). Management of anxiety (panic disorder, with or without agoraphobia, and generalised anxiety disorder) in adults in primary, secondary and community care. 2007; 22.

NICE (National Institute for Health and Clinical Excellence). Osteoarthritis: the care and management of osteoarthritis in adults. 2008; 59.

NICE (National Institute for Health and Clinical Excellence). Depression: The treatment and management of depression in adults. 2009a; 90.

NICE (National Institute for Health and Clinical Excellence). Depression in adults with a chronic physical health problem: Treatment and management. 2009b; 91.

NICE (National Institute for Health and Clinical Excellence). Chronic heart failure. Management of chronic heart failure in adults in primary and secondary care. 2010; 108.

Natvig B, Bruusgaard D, Eriksen W. Localized low back pain and low back pain as part of widespread musculoskeletal pain: two different disorders? A cross-sectional population study. Journal of Rehabilitation Medicine 2001; 33(1): 21-25.

Naylor C, Parsonage M, McDaid D, Knapp M, Fossey M, Galea A. Long-term conditions and mental health: the cost of co-morbidities. London: The King's Fund and Centre for Mental Health; 2012.

Nease DE, Malouin JM. Depression screening: a practical strategy. Journal of Family Practice 2003; 52(2):118-126.

Nesselroade JR. Some implications of the trait state distinction for the study of development across the life span: The case of personality research. In: Baltes PB, Featherman DL, Lerner RM, editors. Life-span development and behavior Hillsdale, NJ: Erlbaum; 1988. p. Vol. 8, 163-189.

Neumann I, Wegener G, Homberg J, Cohen H, Slattery D, Zohar J, Olivier JD, Mathé AA. Animal models of depression and anxiety: What do they tell us about human condition? Progress in Neuro-Psychopharmacology and Biological Psychiatry 2011; 35(6): 1357-1375.

Newman MG, Zuellig AR, Kachin KE, Constantino MJ, Przeworski A, Erickson T, Przeworski A, Erickson T, Cashman-McGrath L. Preliminary reliability and validity of the generalized anxiety disorder questionnaire-IV: A revised self-report diagnostic measure of generalized anxiety disorder. Behavior Therapy 2002; 33(2): 215-233.

Niti M, Ng T, Kua EH, Ho RCM, Tan CH. Depression and chronic medical illnesses in Asian older adults: the role of subjective health and functional status. International journal of Geriatric Psychiatry 2007; 22(11): 1087-1094.

Norton S, Sacker A, Young A, Done J. Distinct psychological distress trajectories in rheumatoid arthritis: Findings from an inception cohort. Journal of Psychosomatic Research 2011; 71:5.

Nour K, Laforest S, Gignac M, Gauvin L. Appreciating the predicament of housebound older adults with arthritis: portrait of a population. Canadian Journal on Aging 2005; 24(1): 57-69.

Novy D, Nelson D, Berry L, Averill P. What does the Beck Depression Inventory measure in chronic pain? A reappraisal. Pain 1995; 61: 261-270.

Nüesch E, Rutjes AW, Husni E, Welch V, Jüni P. Oral or transdermal opioids for osteoarthritis of the knee or hip. Cochrane Database of Systematic Reviews 2009; 4: CD003115.

Nunnally JC. Psychometric theory. 2nd ed. New York: McGraw-Hill; 1978.

Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. Structural Equation Modeling 2007; 14(4): 535-569.

O'Brien, Robert M. "A Caution Regarding Rules of Thumb for Variance Inflation Factors." Quality and Quantity 2007; 41(5): 673-690.

Office for National Statistics. National Statistics Socio-economic Classification 2010. London: Stationary Office; 2010.

Oladinni O. A survey of inner London general practitioners' attitudes towards depression. Primary Care Psychiatry 2002; 8(3):95-98.

Olino TM, Klein DN, Lewinsohn PM, Rohde P, Seeley JR. Latent trajectory classes of depressive and anxiety disorders from adolescence to adulthood: descriptions of classes and associations with risk factors. Comprehensive Psychiatry 2010; 51(3): 224-235.

O'Reilly SC, Jones A, Muir KR, Doherty M. Quadriceps weakness in knee osteoarthritis: the effect on pain and disability. Annals of the Rheumatic Diseases 1998; 57(10): 588-594.

Pai M, McCulloch M, Gorman JD, Pai N, Enanoria W, Kennedy G, Tharyan P, Colford JM Jr. Systematic reviews and meta-analyses: an illustrated, step-by-step guide. National Medical Journal of India 2004; 17(2): 86-95.

Pallant JF, Bailey CM. Assessment of the structure of the Hospital Anxiety and Depression Scale in musculoskeletal patients. Health and quality of life outcomes 2005; 3(1): 82.

Patsopoulos NA, Evangelou E, Ioannidis JPA. Sensitivity of between-study heterogeneity in meta-analysis: proposed metrics and empirical evaluation. International journal of epidemiology 2008; 37(5):1148-1157.

Patten SB, Williams JVA, Wang JL. Mental disorders in a population sample with musculoskeletal disorders. BMC Musculoskeletal Disorders 2006; 7: 37.

Pearson Education, Inc. Beck Depression Inventory–II (BDI–II). 2012; Available at: http://www.pearsonassessments.com/HAIWEB/Cultures/en-us/Productdetail.htm?Pid=015-8018-370. Accessed 20/01/2012.

Peat G, Greig J, Wood L, Wilkie R, Thomas E, Croft P. Diagnostic discordance: we cannot agree when to call knee pain 'osteoarthritis'. Family Practice 2005; 22(1):96-102.

Peat G, Thomas E, Handy J, Wood L, Dziedzic K, Myers H, Wilkie R, Duncan R, Hay E, Hill J, Lacey R, Croft P. The Knee Clinical Assessment Study-CAS(K). A prospective study of knee pain and knee osteoarthritis in the general population: baseline recruitment and retention at 18 months. BMC Musculoskeletal Disorders 2006a; 7: 30.

Peat G, Thomas E, Duncan R, Wood L, Hay E, Croft P. Clinical classification criteria for knee osteoarthritis: performance in the general population and primary care. Annals of the Rheumatic Diseases 2006b; 65(10): 1363-1367.

Pedersen SS, Smith ORF, De Vries J, Appels A, Denollet J. Course of anxiety symptoms over an 18-month period in exhausted patients post percutaneous coronary intervention. Psychosomatic Medicine 2008; 70(3): 349-355.

Peng VY. Latent trajectories modelling of developmental profiles of multimorbidity over time in primary care. Keele University: Arthritis Research UK Primary Care Centre; 2011.

Penninx BWJH, Nolen WA, Lamers F, Zitman FG, Smit JH, Spinhoven P, Cuijpers P, Jong PJ, de Marwijk HWJ, van Meer K, van der Verhaak P. Two-year course of depressive and anxiety disorders: Results from the Netherlands Study of Depression and Anxiety (NESDA). Journal of Affective disorders 2011; 133(1-2): 76-85.

Pfizer, Inc. Patient Health questionnaire (PHQ) screeners. 2012; Available at: http://www.phqscreeners.com/overview.aspx?Screener=02_PHQ-9. Accessed 20/01/ 2012.

Phelan E, Williams B, Meeker K, Bonn K, Frederick J, LoGerfo J, Snowden M. A study of the diagnostic accuracy of the PHQ-9 in primary care elderly. BMC family practice 2010; 11(1): 63.

Pignone MP, Gaynes BN, Rushton JL, Burchell CM, Orleans CT, Mulrow CD, Lohr KN. Screening for depression in adults: A summary of the evidence. Annals of Internal Medicine 2002; 136(10): 765-776.

Pilgrim D, Bentall R. The medicalisation of misery: A critical realist analysis of the concept of depression. Journal of Mental Health 1999; 8(3): 261-274.

Pincus T, Callahan LF. Depression scales in Rheumatoid Arthritis. criterion contamination in interpretation of patient responses. Patient Education and Counseling 1993; 20:133-143.

Pincus T, Vogel S, Field AP. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. Spine 2002; 27(5): E109-120.

Pincus T, Vogel S, Burton AK, Santos R, Field AP. Fear avoidance and prognosis in back pain: A systematic review and synthesis of current evidence. Arthritis and Rheumatism 2006; 54(12): 3999-4010.

Pinto-Meza A, Serrano-Blanco A, Peñarrubia MT, Blanco E, Haro JM. Assessing Depression in Primary Care with the PHQ-9: Can It Be Carried Out over the Telephone? Journal of General Internal Medicine 2005; 20(8): 738-742.

Polanczyk G, de Lima M, Horta B, Biederman J, Rohde L. The worldwide prevalence of ADHD: a systematic review and metaregression analysis. American Journal of Psychiatry 2007; 164(6): 942-948.

Pollock K, Grime J. GPs' perspectives on managing time in consultations with patients suffering from depression: a qualitative study. Family Practice 2003; 20(3): 262-269.

Polsky D, Doshi JA, Marcus S, Oslin D, Rothbard A, Thomas N, Thompson CL. Long-term risk for depression symptoms after a medical diagnosis. Archives of Internal Medicine 2005; 165(11): 1260-1266.

Poole H, Bramwell R, Murphy P. The utility of the Beck Depression Inventory Fast Screen (BDI-FS) in a pain clinic population. European Journal of Pain 2009a; 13(8): 865-869.

Poole H, White S, Blake C, Murphy P, Bramwell R. Depression in chronic pain patients: Prevalence and measurement. Pain Practice 2009b; 9(3):173-180.

Porcheret M, Hughes R, Evans D, Jordan K, Whitehurst T, Ogden H, Croft P. Data quality of general practice electronic health records: the impact of a program of assessments, feedback, and training. Journal of the American Medical Informatics Association 2004; 11(1): 78-86.

Porta MS. A dictionary of epidemiology. 5th ed. New York: Oxford University Press, Inc; 2008.

Prince MJ, Harwood R, Thomas A, Mann A. A prospective population-based cohort study of the effects of disablement and social milieu on the onset and maintenance of late-life depression. The Gospel Oak Project VII. Psychological Medicine 1998; 28(2): 337-350.

Rait G, Walters K, Griffin M, Buszewicz M, Petersen I, Nazareth I. Recent trends in the incidence of recorded depression in primary care. British Journal of Psychiatry 2009; 195 (6): 520-524.

RCGP (Royal College of General Practice). Weekly Returns Service Annual Prevalence Report 2006. Birmingham: Birmingham Research Unit; 2006.

RCGP (Royal College of General Practice). Weekly Returns Service Annual Prevalence Report 2007. Birmingham: Birmingham Research Unit; 2007.

Regier DA, Kaelber CT, Rae DS, Farmer ME, Knauper B, Kessler RC, Norquist GS. Limitations of diagnostic criteria and assessment instruments for mental disorders: implications for research and policy. Archives of General Psychiatry 1998; 55(2): 109.

Reilingh ML, Kuijpers T, Tanja-Harfterkamp A, van der Windt DA. Course and prognosis of shoulder symptoms in general practice. Rheumatology 2008; 47(5): 724-730.

Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. Journla of Clinical Epidmiology 2008; 61 (2): 102-109.

Rindskopf D. Trends in categorical data analyses. New, Semi-new, and Recycled ideas. Latent class models. In: Kaplan D, editor. The SAGE handbook of quantitative methodology for the social sciences. Thousand Oakes: Sage Publications; 2004. p. 137-149.

Rogers, Carl May, Dianne Oliver, A. Experiencing depression, experiencing the depressed: the separate worlds of patients and doctors. Journal of Mental Health 2001; 10(3): 317-333.

Rosemann T, Wensing M, Joest K, Backenstrass M, Mahler C, Szecsenyi J. Problems and needs for improving primary care of osteoarthritis patients: the views of patients, general practitioners and practice nurses. BMC Musculoskeletal Disorders 2006; 7(1):48.

Rosemann T, Backenstrass M, Joest K, Rosemann A, Szecsenyi J, Laux G. Predictors of depression in a sample of 1,021 primary care patients with osteoarthritis. Arthritis Care and Research 2007; 57(3): 415-422.

Rosenthal R, DiMatteo MR. Meta-analysis: Recent developments in quantitative methods for literature reviews. Annual Review of Psychology 2001; 52(1): 59-82.

Rosenstiel AK , Keefe FJ.The use of coping strategies in low-back pain patients: relationship to patient characteristics and current adjustment. Pain 1983; 17: 33-40.

Ross L. Screening for mental health problems and memory impairment for people with long term physical health conditions. North East Mental Health Development Unit. 2010; Available at: http://www.nemhdu.org.uk/silo/files/ltcs-screening-tool-for-mental-health-problems--memory-loss.pdf. Accessed 28/02/ 2011.

Roy-Byrne P, Veitengruber JP, Bystritsky A, dlund MJ, Sullivan G, Craske MG, Welch SS, Rose R, Stein MB. Brief intervention for anxiety in primary care patients. Journal of the American Board of Family Medicine 2009; 22(2):175-186.

Roy-Byrne P, Craske MG, Sullivan G, Rose RD, Edlund MJ, Lang AJ, Bystritsky A, Welch SS, Chavira DA, Golinelli D, Campbell-Sills L, Sherbourne CD, Stein MB. Delivery of evidence-based treatment for multiple anxiety disorders in primary care. JAMA 2010; 303(19): 1921-1928.

Roy-Byrne PP, Katon W, Cowley DS, Russo JE, Cohen E, Michelson E, Parrot T. Panic disorder in primary care: biopsychosocial differences between recognized and unrecognized patients. General Hospital Psychiatry 2000; 22(6): 405-411.

Ruiz MA, Zamorano E, García-Campayo J, Pardo A, Freire O, Rejas J. Validity of the GAD-7 scale as an outcome measure of disability in patients with generalized anxiety disorders in primary care. Journal of Affective Disorders 2011; 128(3): 277-286.

Rutledge T, Reis VA, Linke SE, Greenberg BH, Mills PJ. Depression in Heart Failure: A Meta-Analytic Review of Prevalence, Intervention Effects, and Associations With Clinical Outcomes. Journal of the American College of Cardiology 2006; 48(8):1527-1537.

Sacks JJ, Luo YH, Helmick CG. Prevalence of specific types of arthritis and other rheumatic conditions in the ambulatory health care system in the United States, 2001–2005. Arthritis Care and Research 2010; 62(4): 460-464.

Sahlström A, Johnell O, Redlund-Johnell I. The natural course of arthrosis of the knee. Clinical Orthopaedics and Related Research 1997; 340: 152-157.

Salaffi F, Cavalieri F, Nolli M, Ferraccioli G. Analysis of disability in knee osteoarthritis. Relationship with age and psychological variables but not with radiographic score. Journal of Rheumatology 1991; 18(10): 1581-1586.

Sale JEM, Gignac M, Hawker G. The relationship between disease symptoms, life events, coping and treatment, and depression among older adults with osteoarthritis. Journal of Rheumatology 2008; 35(2): 335-342.

Sarason IG, Sarason BR. Abnormal Psychology: The Problem of Maladaptive Behavior. 11th ed. India: Prentice Hall; 2004.

Sartorius N, Üstün TB, Lecrubier Y, Wittchen HU. Depression comorbid with anxiety: Results from the WHO study on" Psychological disorders in primary health care." British Journal of Psychiatry. Supplement 1996; Jun (30): 38-43.

Schoevers R, Beekman A, Deeg D, Hooijer C, Jonker C, Van Tilburg W. The natural history of late-life depression: results from the Amsterdam Study of the Elderly (AMSTEL). Journal of Affective Disorders 2003; 76(1-3): 5-14.

Schouten J, Valkenburg H. Classification criteria: methodological considerations and results from a 12 year following study in the general population. Journal of Rheumatology. Supplement 1995; 43: 44-45.

Schram MT, Frijters D, van de Lisdonk EH, Ploemacher J, de Craen AJM, de Waal MWM, de Rooij FJ, van Heeringa J, Hofman A, Deeg DJH, Schellevis FG. Setting and registry characteristics affect the prevalence and nature of multimorbidity in the elderly. Journal of Clinical Epidemiology 2008; 61(11): 1104-1112.

Schuurmans J, Comijs H, Beekman A, Beurs E, Deeg D, Emmelkamp P, van Dyck R.The outcome of anxiety disorders in older people at 6-year follow-up: results from the Longitudinal Aging Study Amsterdam. Acta Psychiatrica Scandinavica 2005; 111(6): 420-428.

Schwarz G. Estimating the dimension of a model. Annals of Statistics 1978; 6(2): 461-464.

Schwarz A. Diagnostic Test Calulator. 2002; Available at: http://araw.mede.uic.edu/cgi-bin/testcalc.pl?DT=0&Dt=0&dT=0&dt=0&2x2= Compute. Accessed 12/06/2012.

Schwarzer AC, Aprill C, Derby R, Fortin J, Kine G, Bogduk N. Clinical features of patients with pain stemming from the lumbar zygapophyseal joints. Is the lumbar facet syndrome a clinical entity? Spine 1994; 10: 1132–1137.

Scopaz KA, Piva SR, Wisniewski S, Fitzgerald GK. Relationships of fear, anxiety, and depression with physical function in patients with knee osteoarthritis. Archives of Physical Medicine and Rehabilitation 2009; 90(11):1866-1873.

Scudds RJ, Robertson JM. Pain factors associated with physical disability in a sample of community-dwelling senior citizens. Journals Of Gerontology. Series A, Biological Sciences and Medical Sciences 2000 07; 55(7): M393-399.

Seligman MEP, Walker EF, Rosenhan DL. Abnormal psychology. 4th ed. New York: W.W. Norton & Company, Inc; 2001.

Shapiro S. Meta-analysis/Shmeta-analysis. American Journal of Epidemiology 1994; 140: 771-778.

Sharif M, Kirwan JR, Elson CJ, Granell R, Clarke S. Suggestion of nonlinear or phasic progression of knee osteoarthritis based on measurements of serum cartilage oligomeric matrix protein levels over five years. Arthritis and Rheumatism 2004; 50(8): 2479-2488.

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. Journal of Clinical Psychiatry 1998; 59: 22-33.

Sheskin D. Handbook of parametric and nonparametric statistical procedures. United States: Taylor and Francis Ltd; 2007.

SIGN (Scottish Intercollegiate Guidelines NHS). Management of diabetes. 2010; 116.

Sikorski C, Luppa M, König HH, van den Bussche H, Riedel-Heller SG. Does GP training in depression care affect patient outcome?-A systematic review and meta-analysis. BMC health services research 2012; 12(1):10.

Simon GE, VonKorff M. Recognition, management, and outcomes of depression in primary care. Archives of Family Medicine 1995; 4(2): 99.

Simpson A, Richards D, Gask L, Hennessy S, Escott D. Patients' experiences of receiving collaborative care for the treatment of depression in the UK: a qualitative investigation. Mental Health in Family Medicine 2008; 5(2): 95.

Singleton R, Straits B. Approaches to Social Research. 4th ed. New York: Oxford University Press; 2005.

Slade T, Andrews G. DSM-IV and ICD-10 generalized anxiety disorder: discrepant diagnoses and associated disability. Social Psychiatry and Psychiatric Epidemiology 2001; 36(1): 45-51.

Smarr KL, Keefer AL. Measures of depression and depression symptoms: Beck Depression Inventory-II (BDI-II), Center for Epidemiologic Studies Depression Scale (CES-D), Geriatric Depression Scale (GDS), Hospital Anxiety and Depression Scale (HADS), and Patient Health Questionnaire-9 (PHQ-9). Arthritis Care and Research 2011; 63(11): 454-466

Smith BH, Penny KI, Purves AM, Munro C, Wilson B, Grimshaw J, Chambers WA, Smith WC. The Chronic Pain Grade questionnaire: validation and reliability in postal research. Pain 1997; 71(2): 141-147.

Snowden M, Steinman L, Frederick J, Wilson N. Screening for depression in older adults: recommended instruments and considerations for community-based practice. Clinical Geriatrics 2009; September: 26-32.

Solomon DA, Keller MB, Leon AC, Mueller TI, Shea MT, Warshaw M, Maser JD, Coryell W, Endicott J. Recovery from major depression: a 10-year prospective follow-up across multiple episodes. Arch Gen Psychiatry 1997; 54(11):1001-1006.

Song F, Sheldon TA, Sutton AJ, Abrams KR, Jones DR. Methods for exploring heterogeneity in meta-analysis. Evaluation and The Health Professions 2001; 24(2): 126-151.

Spielberger CD, Rickman RL. Assessment of State and Trait Anxiety. In: Andreoli V, Cassono E, G., Kielholz P, Pancher P, Racagni G, editors. Anxiety: Psychobiological and clinical perspectives Washington: Taylor and Francis; 1991. p. 69-83.

Spies-Dorgelo MN, van der Windt DAWM, van der Horst HE, Prins APA, Stalman WAB. Hand and wrist problems in general practice—patient characteristics and factors related to symptom severity. Rheumatology 2007; 46(11): 1723-1728.

Spitzer RL, Kroenke K, Williams JBW. Validation and Utility of a Self-report Version of PRIME-MD. The PHQ Primary Care Study. JAMA 1999; 282(18): 1737-1744.

Spitzer RL, Kroenke K, Williams JBW, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Archives of Internal Medicine 2006; 166(10): 1092-1097.

Stein MB, Sherbourne CD, Craske MG, Means-Christensen A, Bystritsky A, Katon W, Sullivan G, Roy-Byrne PP. Quality of care for primary care patients with anxiety disorders. American Journal of Psychiatry 2004; 161(12): 2230-2237.

Stengel D, Bauwens K, Sehouli J, Ekkernkamp A, Porzsolt F. Original Paper: A likelihood ratio approach to meta-analysis of diagnostic studies. Journal of Medical Screening 2003; 10(1): 47-51.

Sterne JAC, Bradburn MJ, Egger M. Meta-analysis in STATA. In: Egger M, Smith GD, Altman DG, editor. Systematic Reviews in Health Care: Meta-analysis in context London: BMJ Books; 2003.

Streiner DL. Clinimetrics vs. psychometrics: an unnecessary distinction. Journal of Clinical Epidemiology 2003; 56(12): 1142-1145.

Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. 4th ed. Oxford: Oxford University Press; 2008.

Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology. JAMA 2000; 283(15): 2008-2012.

Schwarzer AC, Aprill C, Derby R, Fortin J, Kine G, Bogduk N. Clinical features of patients with pain stemming from the lumbar zygapophyseal joints. Is the lumbar facet syndrome a clinical entity? Spine 1994; 10: 1132–1137.

Summers MN, Haley WE, Reveille JD, AlarcOan GS. Radiographic assessment and psychologic variables as predictors of pain and functional impairment in osteoarthritis of the knee or hip. Arthritis and Rheumatism 1988; 31(2): 204-209.

Suri P, Morgenroth DC, Kwoh CK, Bean JF, Kalichman L, Hunter DJ. Low back pain and other musculoskeletal pain comorbidities in individuals with symptomatic osteoarthritis of the knee: data from the osteoarthritis initiative. Arthritis Care and Research 2010; 62(12): 1715-1723.

Szoeke C, Dennerstein L, Wluka A, Guthrie J, Taffe J, Clark M, Cicuttini FM. Physician diagnosed arthritis, reported arthritis and radiological non-axial osteoarthritis. Osteoarthritis and Cartilage 2008; 16(7): 846-850.

Tallini A. How should we define health? Health is state of physical, mental, and social wellbeing. BMJ 2011; 343: d5358.

Tallon D, Chard J, Dieppe P. Exploring the priorities of patients with osteoarthritis of the knee. Arthritis Care and Research 2000; 13(5): 312-319.

Tan AL, Grainger AJ, Tanner SF, Shelley DM, Pease C, Emery P, McGonagle D. High-resolution magnetic resonance imaging for the assessment of hand osteoarthritis. Arthritis and Rheumatism 2005; 52(8): 2355-2365.

Telford R, Hutchinson A, Jones R, Rix S, Howe A. Obstacles to effective treatment of depression: a general practice perspective. Family Practice 2002; 19(1): 45-52.

Terluin B, Brouwers E, van Marwijk H, Verhaak P, van der Horst H. Detecting depressive and anxiety disorders in distressed patients in primary care; comparative diagnostic accuracy of the Four-Dimensional Symptom Questionnaire (4DSQ) and the Hospital Anxiety and Depression Scale (HADS). BMC family practice 2009; 10(1): 58.

Terwee CB, Bot SDM, De Boer MR, van der Windt DAWM, Knol DL, Dekker J, Bouter LM, de Vet HCW. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of Clinical Epidemiology 2007; 60(1): 34-42.

Thomas E, Peat G, Harris L, Wilkie R, Croft PR. The prevalence of pain and pain interference in a general population of older adults: cross-sectional findings from the North Staffordshire Osteoarthritis Project (NorStOP). Pain 2004; 110(1-2): 361-368.

Thombs BD, Taillefer SS, Hudson M, Baron M. Depression in patients with systemic sclerosis: A systematic review of the evidence. Arthritis Care and Research 2007; 57(6): 1089-1097.

Thorstensson CA, Gooberman-Hill R, Adamson J, Williams S, Dieppe P. Help-seeking behaviour among people living with chronic hip or knee pain in the community. BMC Musculoskeletal Disorders 2009; 7(10): 153-163.

Timonen M, Liukkonen T. Management of depression in adults. BMJ 2008; 336(7641): 435.

Titterington DM, Smith AFM, Makov UE. Statistical analysis of finite mixture distributions. New York: Wiley; 1985.

Towheed T, Maxwell L, Judd M, Catton M, Hochberg MC, Wells G. Acetaminophen for osteoarthritis. Cochrane Database of Systematic Reviews 2006; (1): CD004257.

Tsai PF. Predictors of distress and depression in elders with arthritic pain. Journal of Advanced Nursing 2005; 51(2):158-165.

Tsuji T, Matsuyama Y, Goto M, Yimin Y, Sato K, Hasegawa Y, Ishiguro N. Knee–spine syndrome: correlation between sacral inclination and patellofemoral joint pain. Journal of Orthopaedic Science 2002;7(5): 519-523.

Tu K, Klein-Geltink J, Mitiku TF, Mihai C, Martin J. De-identification of primary care electronic medical records free-text data in Ontario, Canada. BMC medical informatics and decision making 2010; 10(1):35.

Turk DC, Okifuji A. Detecting depression in chronic pain patients: adequacy of self-reports. Behaviour Research and Therapy 1994; 32(1): 9-16.

Turner JA, Romano JM. Self-report screening measures for depression in chronic pain patients. Journal of Clinical Psychology 1984; 40(4):909-13.

Tylee A, Gastpar M, Lepine J, Mendlewicz J. DEPRES II (Depression Research in European Society II): a patient survey of the symptoms, disability and current management of depression in the community. DEPRES Steering Committee. International Clinical Psychopharmacology 1999; 14(3): 139.

Tylee A, Walters P. Underrecognition of anxiety and mood disorders in primary care: why does the problem exist and what can be done? Journal of Clinical Psychiatry 2007; 68: 27.

UK National Screening Committee. Criteria for appraising the viability, effectiveness and appropriateness of a screening programme. 2009; Available at: www.screening.nhs.uk. Accessed 20/11/2010.

U.S. Preventive Services Task Force (USPSTF). Screening for depression in adults: U.S. Preventive Services Task Force recommendation statement. Annals of Internal Medicine 2009; 151: 784-792.

U.S. Preventive Services Task Force. Screening for Depression: rationale and recommendations. Annals of Internal Medicine 2002; 136(10): 760-764.

Unützer J, Patrick DL, Diehr P, Simon G, Grembowski D, Katon W. Quality adjusted life years in older adults with depression symptoms and chronic medical disorders. International Psychogeriatrics 2000; 12(01): 15-33.

Unützer J, Katon W, Callahan CM, Williams JW, Hunkeler E, Harpole L, Hoffing M, Della Penna RD, Noël PH, Lin EH, Areán PA, Hegel MT, Tang L, Belin TR, Oishi S, Langston C. Depression treatment in a sample of 1,801 depressed older adults in primary care. Journal of the American Geriatrics Society 2003; 51(4): 505-514.

Uthman O. Prevalence and pattern of HIV-related malnutrition among women in sub-Saharan Africa: a meta-analysis of demographic health surveys. BMC Public Health 2008; 8(1): 226-233.

van Beljouw I, Verhaak P, Cuijpers P, van Marwijk H, Penninx B. The course of untreated anxiety and depression, and determinants of poor one-year outcome: a one-year cohort study. BMC Psychiatry 2010; 10(1): 86.

van den Brink RHS, Ormel J, Tiemens BG, Smit A, Jenner JA, van der Meer K, van Os TW. Predictability of the one-year course of depression and generalized anxiety in primary care. General Hospital Psychiatry 2002; 24(3):156-163.

van Dijk GM, Dekker J, Veenhof C, van den Ende CHM. Course of functional status and pain in osteoarthritis of the hip or knee: A systematic review of the literature. Arthritis Care and Research 2006; 55(5): 779-785.

Van Os J, Delespaul P. Toward a world consensus on prevention of schizophrenia. Dialogues in Clinical Neuroscience 2005; 7(1): 53-67.

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(1): 52.

van Ryn M, Burke J. The effect of patient race and socio-economic status on physicians' perceptions of patients. Social Science and Medicine 2000; 50(6): 813-828.

Vartanian TP. Secondary data analysis. New York: Oxford University Press; 2010.

Vermunt JK, Magidson J. Technical guide for Latent GOLD 4.0: Basic and advanced. Belmont Massachusetts: Statistical Innovations, Inc. 2005.

Vermunt JK, Magidson J. Latent class cluster analysis. In: Hagenaars JA, McCutcheon AL, editors. Applied Latent Class Analysis. New York: Cambridge University Press; 2002. p. 89-106.

Viinamäki H, Mäntyselkä P, Kumpusalo E, Halonen P, Takala J, Kauhanen J. Sleep disturbance is associated with mental disorder in patients with hip or knee osteoarthritis. European Journal of Psychiatry 2002; 16(1): 9-13.

Vilalta-Franch J, Garre-Olmo J, López-Pousa S, Turon-Estrada A, Lozano-Gallego M, Hernàndez-Ferràndiz M, Pericot-Nierga I, Garre-Olmo J. Comparison of different clinical diagnostic criteria for depression in Alzheimer disease. American Journal of Geriatric Psychiatry 2006; 14(7): 589.

Vlaeyen JWS, Kole-Snijders AMJ, Rotteveel AM, Ruesink R, Heuts PHTG. The role of fear of movement/(re) injury in pain disability. Journal of Occupational Rehabilitation 1995; 5(4): 235-252.

Volkers AC, Nuyen J, Verhaak PFM, Schellevis FG. The problem of diagnosing major depression in elderly primary care patients. Journal of Affective Disorders 2004; 82(2): 259-263.

Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. Pain 1992; 50(2): 133-149.

Vowles KE, McCracken LM. Acceptance and values-based action in chronic pain: a study of treatment effectiveness and process. Journal of Consulting and Clinical Psychology 2008; 76(3): 397-407.

Watson D, Clark LA, Weber K, Assenheimer JS, Strauss ME, McCormick RA. Testing a tripartite model: II. Exploring the symptom structure of anxiety and depression in student, adult, and patient samples. Journal of Abnormal Psychology 1995; 104(1):15-25.

Watson LC, Pignone MP. Screening accuracy for late-life depression in primary care: a systematic review. Journal of Family Practice 2003; 52(12): 956-964.

Watson M, Brookes S, Faulkner A, Kirwan J. Non-aspirin, non-steroidal anti-inflammatory drugs for treating osteoarthritis of the knee. Cochrane Database of Systematic Reviews 2006; (1): CD000142.

Watts S, Bhutani G, Stout I, Ducker G, Cleator P, McGarry J, Day M. Mental health in older adult recipients of primary care services: is depression the key issue? Identification, treatment and the general practitioner. International Journal of Geriatric Psychiatry 2002; 17(5): 427-437.

Webb P, Bain C. Essential epidemiology: an introduction for students and health professionals. 2nd ed. New York: Cambridge University Press; 2011.

Weisberg RB, Dyck I, Culpepper L, Keller MB. (2007). Psychiatric treatment in primary care patients with anxiety disorders: a comparison of care received from primary care providers and psychiatrists. American Journal of Psychiatry 2007; 164 (2):276-82.

Wells KB, Golding JM, Burnam MA. Affective, substance use, and anxiety disorders in persons with arthritis, diabetes, heart disease, high blood pressure, or chronic lung conditions. General Hospital Psychiatry 1989a; 11(5): 320-327.

Wells KB, Stewart A, Hays RD, Burnam MA, Rogers W, Daniels M, Berry S, Greenfield S, Ware J. The functioning and well-being of depressed patients. JAMA 1989b; 262(7): 914-919.

Wetherell JL, Birchler GD, Ramsdell J, Unützer J. Screening for generalized anxiety disorder in geriatric primary care patients. International Journal of Geriatric Psychiatry 2007; 22(2):115-123.

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 medical research methodology 2003; 3: 25.

Whitley GG. Concept analysis of anxiety. International Journal of Nursing Knowledge 1992; 3(3):107-116.

Whitley GG. Expert validation and differentiation of the nursing diagnoses anxiety and fear. International Journal of Nursing Terminologies and Classifications 1994; 5(4):143-150.

Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression. Journal of General Internal Medicine 1997; 12(7): 439-445.

Widiger TA, Sankis LM. Adult psychopathology: Issues and controversies. Annual Review of Psychology 2000; 51(1): 377-404.

Widiger TA, Samuel DB. Diagnostic categories or dimensions? A question for DSM-V. Journal of Abnormal Psychology 2005; 114(4): 494-454.

Widiger TA, Mullins-Sweatt S 2007. Mental disorders as discrete clinical conditions: Dimensional versus categorical classification. In: S.M. TURNER and M. HERSEN, editors, Adult psychopathology and diagnosis. 3rd ed. New York: Wiley, p. 3-33.

Wilcox S, Brenes GA, Levine D, Sevick MA, Shumaker SA, Craven T. Factors related to sleep disturbance in older adults experiencing knee pain or knee pain with radiographic evidence of knee osteoarthritis. Journal of the American Geriatrics Society 2000; 48(10): 1241-1251.

Wilkie R., Peat G., Thomas E., Croft P. Factors associated with restricted mobility outside the home in community-dwelling adults ages fifty years and older with knee pain: An example of use of the international classification of functioning to investigate participation restriction. Arthritis Care and Research 2007; 57(8): 1381-1389.

Wilkinson M, Barczak P. Psychiatric screening in general practice: comparison of the general health questionnaire and the hospital anxiety depression scale. Journal of the Royal College of General Practitioners 1988; 38(312): 311-313.

Williams JW, Pignone M, Ramirez G, Perez Stellato C. Identifying depression in primary care: a literature synthesis of case-finding instruments. General Hospital Psychiatry 2002; 24(4): 225-237.

Williams DA, Farrell MJ, Cunningham J, Gracely RH, Ambrose K, Cupps T, Mohan N, Clauw D. Knee pain and radiographic osteoarthritis interact in the prediction of levels of self-reported disability. Arthritis & Rheumatism-Arthritis Care and Research 2004; 51(4): 558-561.

Willis BH. Spectrum bias - why clinicians need to be cautious when applying diagnostic test studies. Family Practice 2008; 25(5): 390-396.

Wilson JMG, Jungner G. The Principles and Practice of Screening for Disease. Geneva: WHO; 1968.

Wilson KG, Mikail SF, D'Eon JL, Minns JE. Alternative diagnostic criteria for major depressive disorder in patients with chronic pain. Pain 2001;91(3):227-234.

Wolfe F. Determinants of WOMAC function, pain and stiffness scores: evidence for the role of low back pain, symptom counts, fatigue and depression in osteoarthritis, rheumatoid arthritis and fibromyalgia. Rheumatology 1999; 38(4): 355-361.

Wolitzky-Taylor KB, Castriotta N, Lenze EJ, Stanley MA, Craske MG. Anxiety disorders in older adults: a comprehensive review. Depress and Anxiety 2010; 27(2): 190-211.

Woo J, Ho SC, Lau J, Leung PC. Musculoskeletal complaints and associated consequences in elderly Chinese aged 70 years and over. Journal of Rheumatology 1994 10; 21(10):1927-1931.

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(4): 282-288.

Wood L, Peat G, Thomas E, Duncan R. Knee osteoarthritis in community-dwelling older adults: are there characteristic patterns of pain location? Osteoarthritis and cartilage 2007; 15(6): 615-623.

Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. Bulletin of the World Health Organization 2003; 81(9): 646-656.

World Health Organization. World Health Organization (WHO). International Classification of Diseases: Manual of the International Statistical Classification of Diseases, Injuries and Causes of Death, 6th revision. Geneva: World Health Organization; 1948.

World Health Organization (WHO). The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines. Geneva: World Health Organization; 1992.

World Health Organization (WHO). Mental Health: New Understanding, New Hope. Geneva: World Health Organization; 2001.

Yeung A, Fung F, Yu SC, Vorono S, Ly M, Wu S, Fava M. Validation of the Patient Health Questionnaire-9 for depression screening among Chinese Americans. Comprehensive Psychiatry 2008; 49(2): 211-217.

Yohannes AM, Caton S. Management of depression in older people with osteoarthritis: A systematic review. Aging and Mental Health 2010; 14(6):637-651.

Yohannes AM, Willgoss TG, Baldwin RC, Connolly MJ. Depression and anxiety in chronic heart failure and chronic obstructive pulmonary disease: prevalence, relevance, clinical implications and management principles. International Journal of Geriatric Psychiatry 2010; 25(12): 1209-1221.

Zhang W, Doherty M, Leeb B, Alekseeva L, Arden N, Bijlsma J, Dincer F, Dziedzic K, Hauselmann HJ, Kaklamanis P, Kloppenburg M, Lohmander LS, Maheu E, Martin-Mola E, Pavelka K, Punzi L, Reiter S, Smolen J, Verbruggen G, Watt I. EULAR evidence-based recommendations for the diagnosis of hand osteoarthritis: report of a task force of ESCISIT. Annals of the Rheumatic Diseases 2009; 68(1): 8-17.

Zhang W, Doherty M, Peat G, Bierma-Zeinstra MA, Arden NK, Bresnihan B, Herrero-Beaumont G, Kirschner S, Leeb BF, Lohmander LS, Mazières B, Pavelka K, Punzi L, So AK, Tuncer T, Watt I, Bijlsma JW. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. Annals of the Rheumatic Diseases 2010; 69(3): 483-489.

Zhang Y, Jordan JM. Epidemiology of osteoarthritis. Clinics in Geriatric Medicine 2010; 26(3): 355-369.

Zigmond AS, Snaith R. The hospital anxiety and depression scale. Acta Psychiatrica Scandinavica 1983; 67(6): 361-370.

Appendix A: Conference abstracts

Rzewuska M, Mallen CD, Belcher J, Peat G. The prevalence of comorbid depressive disorders and depression symptoms in adults with osteoarthritis and/or joint pain: systematic review and meta-analyses. Society for Academic Primary Care Annual scientific Meeting, Bristol, England, 6-8 July 2011

Introduction: Recent NICE guidance recommends the management of depressive symptoms in chronic physical conditions. As the commonest chronic condition managed in primary care, and linked with depression, osteoarthritis (OA) can be considered within these recommendations. However, with an estimated 8 million OA sufferers in the UK, the scale of detecting and managing comorbid depression may be considerable. Accurate estimates of the prevalence of comorbid depressive disorders and depressive symptoms in people with OA are needed. We conducted a systematic review and meta-analyses to ascertain these estimates.

Method: Electronic bibliographic databases and reference lists were searched from inception to November 2009. Random effects meta-analyses (for log-transformed estimates) were conducted for observational studies in primary care/general populations reporting the prevalence of depressive disorders and/or symptoms in adults with OA/joint pain. Standardised sensitivity analyses were implemented to estimate pooled prevalence with a pre-specified acceptable level of homogeneity ($I^2 \leq 50\%$). Subgroup meta-analyses and meta-regression analyses were conducted.

Results: 10601 titles were identified and 746 abstracts were screened. 54 studies

described in 88 articles were included. There was substantial between-study heterogeneity in findings (sources included study design, study setting, different case definitions, methods of ascertainment, geographical locations, and populations of interest). Overall pooled prevalence estimates were obtained independently for depressive disorders and depressive symptoms including: major depression (8%, 95% CI 7.1-9.1; 25 estimates; I^2=94.1%, p<0.001); dysthymia (2.6%, 95% CI 2-3.6; 19 estimates; I^2=82.2%, p<0.001); mild-severe depressive symptoms (21.2%, 95% CI 17.7-25.2; 27 estimates; I^2=98.2%, p<0.001); moderate-severe depressive symptoms (14.2%, 95% CI 9.5-20.6; 7 estimates; I^2=97%, p<0.001). The following prevalence estimates were pooled from studies with the acceptable level of homogeneity: major depression 6.6% (95% CI 5.8-7.6, 14 estimates); dysthymia 2.7% (95% CI 2.3-3.2, 18 estimates); mild-severe depressive symptoms 23.3% (95% CI 21.2-24.8, 8 estimates); moderate-severe depressive symptoms (unobtainable). Mild-severe depressive symptoms appeared more common in older age groups, women, and primary care patients [the pooled prevalence for primary care was 27.9% (95% CI 25.5-30.6); 6 estimates; I^2=79%, p<0.001].

Conclusions: Comorbid depressive symptoms, typically minor/subsyndromal, are present in one in four primary care patients with OA. Gender, age and health status differences support person centred detection programs.

Rzewuska M, Mallen CD, Peng VY, Belcher J, Peat G. A latent class growth analysis of anxiety symptoms in a longitudinal cohort of primary care patients with symptomatic osteoarthritis. Division of Health Psychology Annual Conference, Southampton, England, 14-15 September 2011

Background: Symptomatic osteoarthritis (OA) is a painful joint disorder. Anxiety symptoms affect 45% of adults with OA and can deteriorate their functioning. An understanding of the natural course of anxiety over time in OA is required to inform GPs, allocate healthcare resources and improve patients' awareness. This study identifies patterns in anxiety development and describes their characteristics among primary care patients with OA.

Methods: Participants were older UK consecutive general practice attendees, with symptomatic OA. Self-complete questionnaires, containing measures of anxiety and depressive symptoms, age, gender, pain status, coping and social status were mailed within 1 week post-consultation and at 3, 6, 12 months. A Latent Class Growth Analysis was used to identify classes of anxiety development over time. Associations between baseline characteristics and class membership were examined.

Results: A 4-class model emerged for 293 participants with complete anxiety data: normal (41.3%); persistent (29.7%); transient (19.1%) and relapsing (9.9%) anxiety patterns. Ineffective coping strategies, poorer pain status and limited physical/economic resources were associated with maladaptive anxiety responses. Age, gender and factors related to social interactions indicate little impacts on progress of anxiety in OA. Examination of risk factors is to be finalised.

Discussion: The identified longitudinal trajectories of anxiety support differences in adaptation to OA. The course of adaptation appears to be related to characteristics of OA as a stressor, effectiveness of coping and availability of fundamental resources. In practice this study clarifies which of primary care attendees with OA should be targeted for identification of anxiety.

Rzewuska M, Mallen CD, Strauss VY, Belcher J, Peat G. The Course of Comorbid Anxiety Symptoms in Patients Presenting to General Practice with Symptomatic Osteoarthritis: A Latent Class Growth Analysis. Rheumatology 2012; 51: 110.

Background: Concurrent elevated anxiety symptoms are common in people with symptomatic OA and contribute to levels of disability. Yet it is unclear how often anxiety symptoms present at the time of seeking formal healthcare for OA represent persistent states of anxiety and what factors are associated with different anxiety symptom trajectories. An understanding of the natural course of anxiety symptoms in patients with OA is required to inform clinicians, allocate healthcare resources and improve patients' awareness.

Methods: Participants were older adults consulting general practice with symptomatic OA. Self-completion questionnaires, containing measures of anxiety and depressive symptoms, age, gender, pain status, coping and social status were mailed within 1 week of the consultation and at 3, 6, 12 months. A person-centred approach applying Latent Class Growth Analysis (LCGA) was used to identify clusters of anxiety symptoms, which were ascertained with cut-off score ≥8 on the Hospital Anxiety and Depression Scale anxiety subscale. Associations between baseline characteristics and cluster membership were examined using multinomial logistic regression (entry probability $p<0.10$).

Results: A 4-cluster LCGA anxiety model was supported in 293 participants with complete anxiety data. Clusters were: no anxiety (41.3%), persistent (29.7%), unstable (19.1%) and progressive (9.9%) anxiety. Catastrophising, coping by increased behavioural activities, pain extent and interference with work, occupational class and perceived lack of instrumental support were differently

associated with four anxiety clusters (Table). Age, gender, other coping strategies and factors related to social interactions showed no significant effects on anxiety trajectories.

Conclusions: Sixty percent of patients with OA have reported non-normal anxiety levels over 12 months. In addition, an estimated 60% of patients with symptomatic OA presenting to general practice with concurrent anxiety symptoms will experience persistent anxiety for at least 12 months. Odds ratios suggest that coping by catastrophising (Adj. OR=4.25, 95% CI 0.24-0.83) and pain extent (Adj. OR=1.09, 95% CI 1.04-1.15) are most prominent factors associated with the persistent anxiety trajectory.

Table. Backward elimination multinomial logistic regression analyses of baseline covariates and trajectories of anxiety.

	'Persistent' vs. 'No anxiety'*	'Unstable' vs. 'No anxiety'*	'Progressive' vs. 'No anxiety'*
Covariates	Adj. OR (95%CI)	Adj. OR (95%CI)	Adj. OR (95%CI)
Coping by catastrophising	4.25 (1.99, 9.06)	4.04 (1.83, 8.95)	0.93 (0.28, 3.08)
Coping by increased behavioural activities	0.45 (0.24, 0.83)	0.51 (0.27, 0.97)	0.66 (0.27, 1.61)
Lack of instrumental support	5.49 (1.52, 19.81)	1.54 (0.31, 7.60)	3.35 (0.54, 20.82)
Manual/routine occupational class	1.47 (0.76, 2.81)	1.47 (0.75, 2.88)	4.00 (1.60, 9.97)
Pain extent (total number of pain sites 0-44)	1.09 (1.04, 1.15)	1.04 (0.98, 1.09)	1.08 (1.01, 1.15)
Pain interference with work (0-10)	1.18 (1.05, 1.33)	1.12 (0.99, 1.26)	1.29 (1.08, 1.54)

Note:*a reference category

Appendix B: Prevalence of depression and anxiety in adults with osteoarthritis/joint pain: a systematic review and meta-analyses

This appendix supports analyses described in chapter two

B.1 SEARCH STRATEGY

Box B.1.1 Terms used to search six electronic databases.

EMBASE:

Osteoarthritis/joint pain

Keywords: hand, knee, ankle, foot, shoulder, elbow, wrist, hip, pain adj8 psych*, musculoskelet*, spondylosis, osteoarthos*, joint adj2 stiff*, joint adj2 pain, joint adj2 diseas*, arthrit*, Osteoarthrit*

Thesaurus: exp spondylosis, exp cartilage, exp cartilage degeneration, exp elbow, exp elbow disease, exp hand, hand joint, exp hip, exp hip pain, hip osteoarthritis, exp hip disease, exp knee, exp knee disease, knee osteoarthritis, knee pain, exp arthritis, exp arthralgia, exp shoulder, exp shoulder pain, exp shoulder disease, exp wrist, exp wrist disease, exp ankle, exp ankle pain, exp foot, exp foot disease, exp foot pain, exp joint, exp joint degeneration, exp joint stiffness, exp finger joint, exp interphalangeal joint, exp carpometacarpal joint, exp carpal joint, exp atlantoocipital joint, exp atlantoaxial joint, exp acromioclavicular joint, exp toe joint, exp radioulnar joint, exp sacroiliac joint, exp subtalar joint, exp sternocostal joint, exp sternoclavicular joint, exp tarsal joint, exp tarsometatarsal joint, exp temporomandibular joint, exp zygapophyseal joint, exp metacarpophalangeal joint, exp metatarsophalangeal joint, proximal interhalangeal joint, exp patellofemoral joint

Depression/anxiety

Keywords: outcome***,** predictor*, anxiety*, coping, adapt*, anxious adj2 person, anxiety adj2 symptom*, depress* adj2 symptom*, depress*, affect adj2 symptom*, psychol* adj2 distress, antidepressant*, emotional adj2 distress, mood, HADS, hospital and anxiety and depression and scale, CES-D, centre and for and epidemiologic and studies and depression and scale, BDI, beck and depression and inventory, GDS, geriatric and depression and scale, STAI, state and trait and anxiety and inventory, PHQ, patient and health and questionnaire

Box B.1.1 cont. Terms used to search six electronic databases.

Thesaurus: exp mental disease, exp adaptation, exp coping behaviour, exp anxiety, exp anxiety disorder, exp beck anxiety inventory, exp hamilton anxiety scale, exp hospital anxiety and depression scale, exp mixed anxiety and depression, exp self-rating anxiety scale, exp state trait anxiety inventory, exp self-rating depression scale, exp depression, exp rating scale, exp center for epidemiological studies depression scale, exp depression, exp depression inventory, exp geriatric depression scale, exp major depression, exp psychological rating scale, exp mental stress, exp affect, exp emotional, exp mood disorder, exp lifestyle, exp mental health, exp personality test, exp clinical psychology, exp medical psychology, exp risk assessment, exp risk factor, exp quality of life

Setting

Keywords: population* adj2 survey*, primary adj2 care, general adj2 practic*, family adj2 practic*, family adj2 medicine, general adj2 practition*, family adj2 practition*, community adj2 dwell*, general adj2 popul*

Thesaurus: exp general practice, exp primary health care, exp general practitioner

Study design

Keywords: comorbid*, co-occurrence**,** prevalence*, frequency, observational adj2 stud*, cross-sectional adj2 stud*, prospective adj2 stud*

Thesaurus: exp comorbidity, exp incidence, exp prevalence, exp cross-sectional study, exp epidemiology, exp case control study, exp cohort analysis, exp observational study, exp prospective study, exp sickness impact profile, exp health status, exp questionnaire, exp health survey

MEDLINE:

Osteoarthritis/joint pain

Keywords: hand, knee, ankle, foot, shoulder, elbow, wrist, hip, musculoskelet*, spondylosis, osteoarthos*, joint N2 stiff*, joint N2 pain, joint N8 diseas*, arthrit*, osteoarthrit*

Thesaurus: MH cartilage, MH cartilage diseases+, MH elbow, MH elbow joint, MH hand+, MH hand joints+, MH hip, MH hip joint, MH knee, MH knee joint+, MH osteoarthritis+, MH osteoarthritis spine, MH osteoarthritis hip, MH osteoarthritis knee, MH muskuloskuloskeletal diseases/PX, MH arthralgia+, MH shoulder, MH shoulder joint, MH shoulder pain, MH wrist, MH wrist joint, MH ankle, MH ankle joint, MH foot+, MH joint diseases+, MH pain/PX, MH spondylosis+, MH arthritis/EP/PX

Box B.1.1 cont. Terms used to search six electronic databases.

Depression/anxiety

Keywords: outcome***,** predictor*, anxiet*, coping, adapt*, anxious N2 person, anxiety N2 symptom*, depress* N2 symptom*, depress*, affect N2 symptom*, psychol* N2 distress, antidepressant*, emotional N2 distress, mood, HADS, hospital anxiety depression scale, CES-D, centre for epidemiologic studies depression scale, BDI, beck depression inventory, GDS, geriatric depression scale, STAI, state trait anxiety inventory, PHQ, patient health questionnaire

Thesaurus: MH adaptation psychological+, MH anxiety+, MH anxiety disorders+, MH psychiatric status rating scales+, MH stress psychological+, MH affect+, MH affective symptoms, MH mood disorders+, MH life style+, MH mental health, MH personality inventory+, MH psychology clinical, MH psychology medical, MH risk assessment+, MH risk factors, MH quality of life

Setting

Keywords: population* N2 survey*, primary N2 care, general N2 practic*, family N2 practic*, family N2 medicine, general N2 practition*, family N2 practition*, community N2 dwell*, general N2 popul*

Thesaurus: MH family practice, MH primary health care+, MH physicians family

Study design

Keywords: comorbid*, co-occurrence**,** prevalence*, frequency, observational N2 stud*, cohort N2 stud*, cross-sectional N2 stud*, prospective N2 stud*

Thesaurus: MH comorbidity, MH prevalence, MH incidence, MH cross-sectional studies, MH case-control studies+, MH cohort studies+, MH epidemiologic studies+, MH follow-up studies, MH epidemiology+, MH epidemiologic factors+, Mh epidemiologic methods+, MH epidemiologic research design+, MH prospective studies, MH sickness impact profile, MH health status indicators+, MH health surveys+, MH health status+, MH questionnaires+

PSYCINFO:

Osteoarthritis/joint pain

Keywords: hand, knee, ankle, foot, shoulder, elbow, wrist, hip, musculoskelet*, spondylosis, osteoarthos*, joint N2 stiff*, joint N2 pain, joint N8 diseas*, arthrit*, osteoarth*

Thesaurus: MH hip, MH hip joint, MH knee, MH knee joint+, MH cartilage diseases+, MH elbow, MH elbow joint, MH hand+, MH hand joint, MH osteoarthritis+, MH osteoarthritis spine, MH osteoarthritis hip, MH osteoarthritis knee, MH muskuloskletal diseases/PX, MH arthralgia+, MH shoulder, MH shoulder joint, MH shoulder pain, MH wrist, MH wrist joint, MH ankle, MH ankle joint, MH joint diseases+, MH pain/PX, MH spondylosis+, MH arthritis/EP/PX

Box B.1.1 cont. Terms used to search six electronic databases.

Depression/anxiety

Keywords: outcome***,** predictor*, anxiet*, coping, adapt*, anxious N2 person, anxiety N2 symptom*, depress* N2 symptom*, depress*, affect N2 symptom*, psychol* N2 distress, antidepressant*, emotional N2 distress, mood, HADS, hospital anxiety depression scale, CES-D, centre for epidemiologic studies depression scale, BDI, beck depression inventory, GDS, geriatric depression scale, STAI, state trait anxiety inventory, PHQ, patient health questionnaire

Thesaurus: MH adaptation psychological+, MH anxiety+, MH anxiety disorders+, MH psychiatric status rating scales+, MH stress psychological+, MH affect+, MH affective symptoms, MH mood disorders+, MH life style+, MH mental health, MH personality inventory+, MH psychology clinical, MH psychology medical, MH risk assessment+, MH risk factors, MH quality of life

Setting

Keywords: population* N2 survey*, primary N2 care, general N2 practic*, family N2 practic*, family N2 medicine, general N2 practition*, family N2 practition*, community N2 dwell*, general N2 popul*

Thesaurus: MH family practice, MH primary health care+, MH physicians family

Study design

Keywords: comorbid*, co-occurrence**,** prevalence*, frequency, observational N2 stud*, cohort N2 stud*, cross-sectional N2 stud*, prospective N2 stud*

Thesaurus: MH comorbidity, MH prevalence, MH incidence, MH cross-sectional studies, MH case-control studies+, MH cohort studies+, MH epidemiologic studies+, MH follow-up studies, MH epidemiology+, MH epidemiologic factors+, MH epidemiologic methods+, MH epidemiologic research design+, MH prospective studies, MH sickness impact profile, MH health status indicators+, MH health surveys+, MH health status+, MH questionnaires+

CINAHL:

Osteoarthritis/joint pain

Keywords: hand, knee, ankle, foot, shoulder, elbow, wrist, hip, musculoskelet*, spondylosis, osteoarthos*, joint N2 stiff*, joint N2 pain, joint N8 diseas*, arthrit*, osteoarthrit*

Thesaurus: MH hip, MH hip joint, MH knee, MH knee joint+, MH cartilage diseases+, MH elbow, MH elbow joint, MH hand+, MH foot, MH hand joint, MH osteoarthritis+, MH osteoarthritis spine, MH osteoarthritis hip, MH osteoarthritis knee, MH muskuloskletal diseases/PX, MH arthralgia+, MH shoulder, MH shoulder joint, MH shoulder pain, MH wrist, MH wrist joint, MH ankle, MH ankle joint, MH joint diseases+, MH pain/PX, MH spondylosis+, MH arthritis/EP/PX

Box B.1.1 cont. Terms used to search six electronic databases.

Depression/anxiety
Keywords: outcome***,** predictor*, anxiet*, coping, adapt*, anxious N2 person, anxiety N2 symptom*, depress* N2 symptom*, depress*, affect N2 symptom*, psychol* N2 distress, antidepressant*, emotional N2 distress, mood, HADS, hospital anxiety depression scale, CES-D, centre for epidemiologic studies depression scale, BDI, beck depression inventory, GDS, geriatric depression scale, STAI, state trait anxiety inventory, PHQ, patient health questionnaire
Thesaurus: MH adaptation psychological+, MH anxiety+, MH anxiety disorders+, MH psychiatric status rating scales+, MH stress psychological+, MH affect+, MH affective symptoms, MH depression, MH depressive disorder+, MH depressive disorder major+, MH mood disorders+, MH life style+, MH mental health, MH personality inventory+, MH psychology clinical, MH psychology medical, MH risk assessment+, MH risk factors, MH quality of life
Setting
Keywords: population* N2 survey*, primary N2 care, general N2 practic*, family N2 practic*, family N2 medicine, general N2 practition*, family N2 practition*, community N2 dwell*, general N2 popul*
Thesaurus: MH family practice, MH primary health care+, MH physicians family
Study design
Keywords: comorbid*, co-occurrence**,** prevalence*, frequency, observational N2 stud*, cohort N2 stud*, cross-sectional N2 stud*, prospective N2 stud*
Thesaurus: MH comorbidity, MH prevalence, MH incidence, MH cross-sectional studies, MH case-control studies+, MH cohort studies+, MH epidemiologic studies+, MH follow-up studies, MH epidemiology+, MH epidemiologic factors+, MH epidemiologic methods+, MH epidemiologic research design+, MH prospective studies, MH sickness impact profile, MH health status indicators+, MH health surveys+, MH health status+, MH questionnaires+
ISI Web of Knowledge and CSA ILLUMINA
(anxiet*or depress*) and (osteoarthrit*or arthrit* or joints)

Note: Terms belonging to the same categories were linked with 'OR' and categories were linked with 'AND'

B.2 DETAILS OF INCLUDED STUDIES

Table B.2.1 Overview of studies from which data was synthesised.

	Methodological summary						Statistical summary for cases with OA/joint pain		
Study ID [ID number]	**Country/ region**	**Design (response rate)**	**Setting**	**Mean age (SD): range**	**Grouping category (definition used by authors): description of cases ascertainment/time scale(anatomical site)**	**Sample (% F)**	**Mental health problem: assessment tool (range); cut-off score/severity**	**Prevalence rate Of mental health problem in OA/joint pain and compared groups**	**Mean score**
Questionnaires/self-reported anxiety and depression ascertainment									
Allen, 2008 [1]	U.S.A/ Johnston county	Prospective-cohort (BL: 95%)	General population	≥45	Radiographic OA (OA): radiographs graded with the K+L criteria, 1 question on joint pain or stiffens/most days (knee, hip)	With OA: 759 (67.3%) Without OA: 1923 (65%)	Depression: CESD-20 (0-60);≥16/mild or worse	In OA: 13.3% Without OA: 12.7% p>0.05	
Barberger-Gateau,1992 [2]	France/ Gironde, Dordogne	Cross-sectional (BL: 68.9%)	General population	With joint pain: 75: ≥65	Symptomatic OA (joint pain): one question/current (all)	Total: 2792 (60%) Joint pain: 1987	Depression: CESD-20 (0-60); ≥16/mild or worse	In joint pain: 15.8%	
Creamer, 1999 [3]	U.S.A/ Baltimore	Cross-sectional (BL: 88.1%)	General Population	With OA: 64.3 (0.52): ≥ 40	Radiographic OA (OA): radiographs graded with the K+L criteria, pain assessed with the NHANES/ever, for most days for at least month (knee)	OA: 374 (31.5%)	Depression: AIMS-D (0-10); ≥4/moderate or worse	In OA: 3.2%	In OA: Total: 1.25(SD 1.09) F: 1.47 (SD 1.19) M: 1.14 (SD 1.03) p=0.006
Croft, 2005 [4]	UK / North Staffordshire	Cross-sectional (BL: 70%)	Primary Care	Total: 65 (10.0): ≥ 50	Symptomatic OA (joint pain): joint pain assessed with a pain manikin and the WOMAC/ for more than 1 day in the last month (knee, elsewhere- neck, hand lower back, hip, foot, ankle)	Total: 5364 (55%) No pain: 1909 (51%) Knee pain alone: 457 (49%) Knee pain plus 1 pain: 496 (55%) Knee pain plus ≥ 2 pains: 1257(62%)	Anxiety: HADS-A (0-21), above the highest tertile/ moderate or worse Depression: HADS – D (0-21); above the highest tertile/ moderate or worse	No pain:18% In knee pain alone: 23% In plus 1 pain: 32% In plus ≥ 2 pains: 43% No pain:16% In knee pain alone: 19% In plus 1 pain: 29% In plus ≥ 2 pains: 42%	
Dexter, 1994 [5]	U.S.A/ Indianapolis	Cross-sectional	General Population	75.4 (9.45): ≥50	The ACR criteria diagnosed OA (OA): the ACR criteria for OA (hip, knee)	OA: 108 (88%)	Depression: AIMS-D (0-10); ≥ 5/moderate or worse	In OA: 12%	In OA: 2.23 (SD 1.89)

Table B.2.1 cont. Overview of studies from which data was synthesised.

	Methodological summary						Statistical summary for cases with OA/joint pain		
Study ID [ID number]	**Country/region**	**Design (response rate)**	**Setting**	**Mean age (SD): range**	**Grouping category (definition used by authors): description of cases ascertainment/time scale(anatomical site)**	**Sample (%F)**	**Mental health problem: assessment tool (range); cut-off score/severity**	**Prevalence rate Of mental health problem in OA/joint pain and compared groups**	**Mean score**
Dunlop, 2005 [6]	U.S.A/ national	Prospective -cohort (BL: 94%)	General Population	≥54	Arthritis (arthritis): self-reported, physician-diagnosed arthritis/rheumatism/ lifetime or current (all)	Arthritis: 5715(64.19%)	Depression: CESD-5 (0-15)	In arthritis: Total: 40.3% F: 44.2% M: 33.3%	
Figaro, 2005 [7]	U.S.A / New York	Cross-sectional (BL: 85.1%)	General population	71 (8): ≥50	The ACR criteria diagnosed OA (OA) : OA assessed with screening tool consistent with criteria developed by Altman et al. (1986), pain assessed with the WOMAC (knee)	OA: 94 (84%)	Depression: 2 questions on low mood and anhedonia (Whooley & Simon, 2000); yes to one/?	In OA: Current: 22% Lifetime: 14%	
Fisher, 2004 [8]	U.S.A/Texas, California, Arizona, Colorado, New York	Prospective -cohort (BL: 83%)	General Population	72.8 (6.3): ≥65	Arthritis (arthritis): self-reported, physician-diagnosed arthritis /ever (all)	Arthritis: 1084 (70.7%)	Negative affect: the 7-item scale formulated from CESD (0-21): score ≥ 1	In arthritis: 23.2%	In arthritis: 3.9 (SD 4.6)
Hill, Dziedzic, Thomas 2007a [9]	UK /North Staffordshire	Cross-sectional (BL: 78.6%)	Primary Care	Total: 65.4 (9.6): ≥ 50	Self-reported OA (OA): questions on self-reported GP-diagnosed OA (hand)	Total: 2113 (56.1%) OA: 538* *data reported by n=535	Anxiety: HADS-A (0-21); 8-21- mild or worse	In OA: 50%	In OA: 7.90 (SD 4.4)
							Depression: HADS-D (0-21); 8-21- mild or worse	In OA: 26%	In OA: 5.34 (SD 3.7)
Jakobsson, 2006 [10]	Sweden/ national	Cross-sectional (BL: 49%)	General Population	With OA: 83.8 (5.8) Reference group: 82.7 (5.6): ≥75	Self-reported OA (OA): self-reported OA / past 3 months (all)	OA: 168 (66.1%) Reference group: 246 (58.9%)	Depressed mood: Altman, 1999 No, not at all Yes, little Yes, rather much Yes, very much	In OA (in reference group): No, not at all 63.7 % (66.3%) Yes, little 23.8% (18.7%) Yes, rather much 4.8 % (5.7%)	

Table B.2.1 cont. Overview of studies from which data was synthesised.

	Methodological summary						Statistical summary for cases with OA/joint pain		
Study ID [ID number]	**Country/ region**	**Design (response rate)**	**Setting**	**Mean age (SD): range**	**Grouping category (definition used by authors): description of cases ascertainment/time scale(anatomical site)**	**Sample (%F)**	**Mental health problem: assessment tool (range); cut-off score/severity**	**Prevalence rate Of mental health problem in OA/joint pain and compared groups**	**Mean score**
Kramer, 2002 [11]	Netherlands/ Amsterdam	Prospective-cohort (BL: 62.3%)	General Population	68(7.9): ≥ 55	Self-reported OA (OA): self-reported, GP diagnosed OA (all)	Total: 3017 (51.5%) With OA: 1005*(66.9%) Without OA: 2081 *Depression data for n=995	Depression: CESD 20 (0-60); ≥ 16/ mild or worse ≥23/moderate or worse	In OA: Mild or worse: Total: 19.2% F: 21.6% M: 14.4% Moderate or worse: Total: 7.4%	In OA: Total: 9.6 (SD 8.2) M: 8.1 (SD 7.5) F: 10.3 (SD 8.4)
Leveille, 2007 [12]	U.S.A/ East Baltimore	Prospective-cohort (BL : 78%)	General Population	≥ 65	Symptomatic OA (joint pain): self-reported joint pain scored on a 0-10 NRS, the WOMAC/ on most days for at least 1 month in the previous year (knee, hip, foot)	Total: 1002 Analysed: 460 Joint pain: 136	Anxiety: HSCL-4; positive response to at least 2 items; prevalent Depression: GDS-30; ≥ 14/ mild or worse	In no pain: 12.5% In joint pain: 19.1% In other pain: 19% In widespread pain: 28.4 Trend p=0.005 In no pain: 5.2% In joint pain: 9.6% In other pain: 12% In widespread pain: 18.2% Trend p=0.007	
Mallen, 2008 [13]	UK / North Staffordshire	Prospective-cohort (BL crude: 77.2%)	Primary Care	Analysed: 63.1 (10.6): ≥ 50	Symptomatic OA (joint pain): medical record of musculoskeletal pain identified through the EMIS template recorded by GP; pain intensity assed on a 0-10 NRS (all)	Total: 502 (60%) Analysed with joint pain: 428 (59.6%)	Anxiety: HADS-A (0-21); 0-7- no ≥8 mild or worse, ≥11 mild or worse Depression: HADS-D (0-21);0-7-no 8-21-mild 11-14-moderate 15-21-severe 2 questions on low mood and anhedonia, yes to one	In joint pain: No 53.5% Mild or worse: 44.4% Moderate or worse: 21.3% In joint pain: No 63.8% Mild: 21.7% Moderate: 11.4% Severe: 1.9% GP (self-administered) Yes: 18.2% (50.9%) No: 69.4% (48.1%) Unclear: 12.4% (0.9%)	In joint pain:7.31 (SD 4.04) In joint pain:6.1 (SD 3.8)

Table B.2.1 cont. Overview of studies from which data was synthesised.

Methodological summary							Statistical summary for cases with OA/joint pain		
Study ID [ID number]	**Country/ region**	**Design (response rate)**	**Setting**	**Mean age (SD): range**	**Grouping category (definition used by authors): description of cases ascertainment/time scale(anatomical site)**	**Sample (%F)**	**Mental health problem: assessment tool (range); cut-off score/severity**	**Prevalence rate Of mental health problem in OA/joint pain and compared groups**	**Mean score**
Memel, 2000 [14]	UK/Bristol	Cross-sectional (BL: 91%)	Primary Care	71: ≥42	Medical records defined OA (OA): the practices' computerized appointment systems, medical records and diagnoses verified in the paper records (knee, hip)	Eligible: 200 Analysed with OA: 182 (65%)	Anxiety: HADS-A (0-21); 0-7- no 8-10- mild 11-21- moderate or worse	In OA: Self-completed: No: 53.6% Mild: 22% ≥ Moderate: 24.4% GP-assessed: No: 63.7 % Mild: 24.4 % ≥ Moderate: 11.9 %	
							Depression: HADS-D-7 (0-21); 0-7- no 8-10- mild 11-21- moderate or worse	In OA: Self-completed No: 70.8%,Mild: 20.8% ≥ Moderate: 8.3% GP-assessed No: 71.4%,Mild: 22.6% ≥ Moderate: 6.0%	
Muus, 2007 (15)	U.S.A/ national	Cross-sectional	General Population	≥ 55	Arthritis (arthritis): self-reported doctor-diagnosed arthritis (all)	Total: 8305 Arthritis: 3613 (50.2%)	Depression: self-reported doctor-diagnosed	In arthritis : 18.8%	
Niti, 2007 (16)	China/ Singapor e	Cross-sectional (BL: 78.5%)	General Population	≥ 55	Arthritis (arthritis): self-reported, doctor-diagnosed arthritis (all)	Total: 2611 (63.1%), Arthritis: 432	Depression: GDS-15 (0-15); ≥ 5/mild or worse	In arthritis: 16.7%	
Nour, 2005 (17)	Canada/ Montreal	Cross-sectional	General Population	Total: 76.9 (10.5) OA: 79.39(8.25): ≥ 50	Medical records defined OA (OA): medical records on OA (provided by care managers), confirmed by patients, the WOMAC, a 100-mm long VAS for pain intensity (all)	Total: 125 (91.2%) OA: 81	Depression: CES-D-20 (0-60), 0-15- no 16-22- mild 23-29-moderate 30-60- severe	In OA: No: 46.1% Mild: 17.1% Moderate: 17.1% Severe: 19.7%	

Table B.2.1 cont. Overview of studies from which data was synthesised.

	Methodological summary						Statistical summary for cases with OA/joint pain		
Study ID [ID number]	**Country/ region**	**Design (response rate)**	**Setting**	**Mean age (SD): range**	**Grouping category (definition used by authors): description of cases ascertainment/time scale(anatomical site)**	**Sample (%F)**	**Mental health problem: assessment tool (range); cut-off score/severity**	**Prevalence rate Of mental health problem in OA/joint pain and compared groups**	**Mean score**
O'Reilly, 1998 [18]	UK/ Nottingham	Nested case-control (BL: 63.3%)	General Population	Joint pain: 61.3 (10.4) No pain: 60.8 (11.0): ≥ 40	Radiographic OA (OA): 2 questions on joint pain, the WOMAC, radiographs/ever on most days for at least a month and within the last year (knee)	Total: 600 (64%) OA: 300^ Without pain: 300~ ^- anxiety data for n=296; depression data for n=297 ~ anxiety data for n=298; depression data for n=299	Anxiety: HADS-A (0-21); ≥8/mild or worse	In OA: * Total: 39% M: 30 %, F: 45% In without OA:* Total: 24.12% M: 18%, F: 28%	In OA: 7.1(95% CI 6.6 - 7.6) In without OA: 5.4 (95% CI 5.1- 5.8) p<0.005
							Depression: HADS-D (0-21); ≥ 8/mild or worse	In OA*: Total: 19% M: 18%, F: 20% In without OA:* Total: 5.9% M: 4%, F: 7% *read from the graph	In OA: 4.9 (95% CI 4.6- 5.3) In without OA: 2.9 (95%CI 2.6- 3.1) p<0.005
Peat, 2006a [19]	UK/ North Staffordshire	Prospective-cohort (BL: 69.7% FUP 1: 87.7% FUP 2: 42.2% FUP 3: 96.8%)	Primary Care	65 (10): ≥ 50	Symptomatic OA (Joint pain): the Health Survey Questionnaire, and the WOMAC and digital photographs at FUP (knee)	1. BL: 6108 (56%) 2. Reported knee pain in last 12 months: 3106 (59%)* 3. Consented to further contact: 2226 (57%) 4. FUP 1 Respondents to regional pains survey: 1949 (57%) 5. FUP 2 Attended research clinic: 819 (54%) 6. FUP 3 (18 months): 776 (54%) *meta- analysed	Anxiety: HADS-A (0-21), 0-7-no 8-10-mild 11-21- moderate or worse	No: 1 -62%,2 -55%, 3-56%, 4 -57%, 5-62%,6 -63% Mild: 1-25%,2-28%, 3-28%,4-28%, 5-27%,6 -27% Moderate or worse: 1-13%,2-17%, 3-16%,4-15%, 5-11%,6-11%	

Table B.2.1 cont. Overview of studies from which data was synthesised.

	Methodological summary						Statistical summary for cases with OA/joint pain		
Study ID [ID number]	**Country/ region**	**Design (response rate)**	**Setting**	**Mean age (SD): range**	**Grouping category (definition used by authors): description of cases ascertainment/time scale(anatomical site)**	**Sample (%F)**	**Mental health problem: assessment tool (range); cut-off score/severity**	**Prevalence rate Of mental health problem in OA/joint pain and compared groups**	**Mean score**
Peat, 2006a [19]	UK/ North Staffordshire	Prospective-cohort (BL: 69.7% FUP 1: 87.7% FUP 2: 42.2% FUP 3: 96.8%)	Primary Care	65 (10): ≥ 50	Symptomatic OA (Joint pain): the Health Survey Questionnaire, and the WOMAC and digital photographs at FUP (knee)	1. BL: 6108 (56%) 2. Reported knee pain in last 12 months: 3106 (59%)* 3. Consented to further contact: 2226 (57%) 4. FUP 1 Respondents to regional pains survey: 1949 (57%) 5. FUP 2 Attended research clinic: 819 (54%) 6. FUP 3 (18 months): 776 (54%) *meta- analysed	Depression: HADS-D (0-21), 0-7-no 8-10-mild 11-21- moderate or worse	No 1-78%, 2-73%, 3-75%, 4-77%, 5-83%, 6-83% Mild 1-16%, 2-20%, 3-18%,4-17%, 5-13%, 6-12% Moderate or worse 1-6%, 2-7% 3-6%,4-6% 5-4%,6- 5%	
Polsky, 2005 [20]	U.S.A/ national	Prospective-cohort	General Population	OA: 55.9 (3.2): ≥ 51	Arthritis (arthritis): self-reported doctor-diagnosed arthritis/ rheumatism (all)	Total: 8387 (51.8%) Arthritis: 1754 (51%)	Depression: CESD-8 (0-24); ≥ 5	In arthritis: 4.4%	In arthritis: 3.3 (SD 2.5)

Table B.2.1 cont. Overview of studies from which data was synthesised.

	Methodological summary						Statistical summary for cases with OA/joint pain		
Study ID [ID number]	**Country/ region**	**Design (response rate)**	**Setting**	**Mean age (SD): range**	**Grouping category (definition used by authors): description of cases ascertainment/time scale(anatomical site)**	**Sample (%F)**	**Mental health problem: assessment tool (range); cut-off score/severity**	**Prevalence rate Of mental health problem in OA/joint pain and compared groups**	
Roseman n, 2007 [21]	German/ 75 general practices	Cross-sectional (BL: 81.68%)	Primary care	Total: 66.1(15.1): ≥ 18	The ACR criteria defined OA (OA): the ACR criteria and radiographs graded with the K+L (hip, knee)	Total: 1021 (66%) Analysed with OA: 1012	Depression: PHQ-9 (0-27); 1-4- no 5-9- mild 10-14-moderate * 15-19-moderately severe 20-27-severe *a recommended threshold for any clinically relevant symptoms, so equivalent of mild or worse in other scales	In OA: No: Total: 32.9% M: 29.1%, F: 34.9% Mild Total: 38.6% M: 39.8%, F: 38% Moderate Total: 9.1% M: 11.3%, F: 7.9% Mod.-Severe: Total: 14.9% M: 15.1%, F: 14.8% Severe: Total: 4.4% M: 4.7% F: 4.3% Gender diff.: No p=0.11 Mild p=0.06 Moderate p<0.05 Mod. severe p=0.23 Severe p=0.34	In OA: Total: 15.7(SD 4.7) F: 15.33 (SD 4.8) M: 15.95 (SD 4.6)
Sale, 2008 [22]	Canada/ Ontario	Prospective-cohort (BL: 72%)	General Population	75.1 (7.8): ≥ 61	Symptomatic OA (joint pain): difficulty in mobility in the past 3m; joint pain/ stiffness/ swelling; a pain manikin, WOMAC /lasting 6 weeks ≥ in the past 3 months (knee, hip)	OA at BL: 1227 (75.6%)	Depression: CESD-20 (0-60); ≥ 16/mild or worse	In joint pain: 21.3%	In joint pain: 9.4 (SD 8.0)
Schram, 2008 [23]	Netherlan ds/ Leiden	Cross-sectional (BL: 85%)	General Population	≥ 85	Arthritis (arthritis): arthritis identified via GP interviews, medical and pharmacy records (all)	Total: 599 (67%) Joint problems: 171	Depression: GDS-15(0-15); ≥ 4/borderline normal and mild	In arthritis: 25.7%	In arthritis: 2.5 (SD 2.5)

Table B.2.1 cont. Overview of studies from which data was synthesised.

	Methodological summary						Statistical summary for cases with OA/joint pain		
Study ID [ID number]	**Country/ region**	**Design (response rate)**	**Setting**	**Mean age (SD): range**	**Grouping category (definition used by authors): description of cases ascertainment/time scale(anatomical site)**	**Sample (%F)**	**Mental health problem: assessment tool (range); cut-off score/severity**	**Prevalence rate Of mental health problem in OA/joint pain and compared groups**	**Mean score**
Scudds, 2000 [24]	Canada/ Ontario	Cross-sectional (BL: 70.7%)	General population	Musculoskeletal pain 75.8 (5.83): ≥ 65	Symptomatic OA (joint pain): self-reported pain in joint, muscles or bones, a pain manikin/ the past 2 weeks (all)	Total: 884 Musculoskeletal pain : 644 (63.2%)	Depression: CESD-20 (0-60); ≥ 16/mild or worse	In musculoskeletal pain: Total: 25.4% F: 26.5% M: 23.6%	
Szoeke, 2008 [25]	Australia/ Melbourne	Prospective-cohort (BL: 71%)	General Population	59.9 (2.5): ≥ 45	Arthritis (arthritis): self-perceived OA and manikin (all)	Total: 224 (100%) Self-perceived arthritis: 118* *depression data reported for 89	Depression: CESD-10(0-30)	In self-perceived arthritis : 17%* *calculated by the authors of the review	In self-perceived arthritis : 7.08 (SD 4.3) Without self-perceived arthritis : 5.7 (SD 3.5)
Wilcox, 2000 [26]	U.S.A/ Winston-Salem	Prospective cohort (BL: 54.9%)	Primary Care	BL: 71.76 (4.95): ≥ 65	The ACR criteria defined OA (OA): self-reported joint pain, radiographs, the ACR Criteria (knee)	Total OA: 463 Analysed with OA: 429* (52.4%) * avaialble for n=424	Depression: CESD-20 (0-60); ≥ 16/mild or worse	In OA: 17.9%	In OA: 9.59 (SD 7.39)
Wilkie, 2007 [27]	UK/North Staffordshire	Cross-sectional (Stage 1: 71.3%, Stage 2: 88.8%)	Primary Care	65.3 (9.7): ≥ 50	Symptomatic OA (joint pain): joint pain assessed with the WOMAC (knee)	Joint pain: 2252 (58%)	Anxiety: HADS-A (0-21), 0-7-no, 8-10- mild 11-21-moderate or worse Depression: HADS-D (0-21), 0-7-no, 8-10-mild ,11-21-moderate or worse	In joint pain: No: 53.9% Mild: 24.9% ≥ Moderate: 21.3% In joint pain: No: 75.4% Mild: 15.1% ≥ Moderate:9.5%	
Woo, 1994 [28]	China/ Hong-Kong	Cross-sectional (BL: 60%)	General Population	≥70	Symptomatic OA (joint pain): self-reported joint pain / the past 12 months (knee: 65.44%, back-thoracic region- 46.23%, ankle/foot-31.30%, shoulder-33.45%, wrist/finger/ neck-17.5%, elbow-17.67% back- lumpar region-20.15%)	Total: 2023 (51.1%) Pain at various sites: 1166 (60.63%) Pain limiting activities: 600 (67.67%)	Depression: GDS-15 (0-15); ≥ 8/mild or worse	In with pain limiting activities group: 39.8%* *calculated by the authors of the review	

Table B.2.1 cont. Overview of studies from which data was synthesised.

	Methodological summary						Statistical summary for cases with OA/joint pain	
Study ID [ID number]	**Country/ region**	**Design (response rate)**	**Setting**	**Mean age (SD): range**	**Grouping category (definition used by authors): description of cases ascertainment/time scale(anatomical site)**	**Sample (%F)**	**Mental health problem: assessment tool (reference standard), time frame**	**Prevalence rate of mental health problem in OA/joint pain and compared groups**
Clinically diagnosed/clinical interviews data anxiety and depressive disorders								
Dunlop, 2004 [29]	U.S.A/ national	Prospective-cohort (BL: 94%)	General population	≥54	Arthritis (arthritis): self-reported, physician diagnosed arthritis or rheumatism/ lifetime /current (all)	Total at BL: 7825 (52%) Arthritis : 3912	Major depression: WMH WHO-CIDI-SF (DSM-III-R),12 months	In arthritis: 10.88%
Gureje, 2008 [30]	Nigeria/ Yoruba speaking area	Cross-sectional (BL: 74.2%)	General Population	≥65	Arthritis (arthritis): arthritis identified with a checklist of chronic physical and pain conditions (all)	Total: 2152 Arthritis: 1488 (56.6%)	Major depression: WMH WHO-CIDI-III (DSM-IV),12 months	In arthritis: 7.9% (95% CI 6.6-9.4)
He, 2008 [31]	International/ 17 countries in America, Europe, the Middle East, Asia, the South Pacific	Cross-sectional: (BL weighted average: 71% Range: 46%-88%)	General population	Colombia: 36.6 Mexico: 35 U.S.A: 45 Japan: 51 Beijing: 40 Shanghai: 42.9 New Zealand: 46 Belgium: 46.9 France: 46.3 Germany 48.2 Italy: 47.7 Netherlands: 45 Spain: 45 Ukraine: 46 Lebanon: 40 Nigeria: 35.8 Israel: 44.4 South Africa:37.1/≥ 18	Arthritis (arthritis): 2 questions adapted from the US Health Interview Survey (NCHS, 1994): 1. Have you experienced arthritis or rheumatism in the prior 12 months 2. Have you ever had arthritis and rheumatism/ period or lifetime (all)	Total: 85088 Analysed: 42697 (Range 47.5%-53.6%) With arthritis: 7842, including: Colombia: 184 Mexico: 229 U.S.A: 1588 Japan: 117 Beijing, PRC: 111 Shanghai, PRC: 114 New Zealand: 1474 Belgium: 227 France: 432 Germany: 151 Italy: 510 Netherlands: 134 Spain: 617 Ukraine: 479 Lebanon: 57 Nigeria: 469 Israel : 496 South Africa: 453	Anxiety disorders: A. Generalized anxiety disorder B. Panic disorder and/or agoraphobia C. Social phobia D. PTSD Depressive disorders: E. Major depression F. Dysthymia: WMH WHO - CIDI (DSM-IV),12 months	In arthritis: Colombia: A. 1% B. 3.2% C. 3.7% D. 0.6% E. 9.3% F. 1.6% Mexico: A. 1.4% B. 2.9% C. 5.3% D. 0.3% E.10.2% F. 2% U.S.A: A. 5.9% B. 5.0% C.7.5% D. 5.2% E. 9.3 % F.3.6% Japan: A. 2.5% B. 0.9% C. 0% D. 0.1% E. 2.2 % F. 0.1% Beijing: A. 2.3% B. 0.3% C. 0% D. 0% E. 3.9 % F. 1.1% Shanghai: A. 3.6% B. 0% C. 0% D. 0.8% E. 5.7 % F. 0.3% New Zealand: A. 3.9 % B. 2.6% C. 4.7% D. 4 % E. 6.2 % F. 2.4% Belgium: A. 1.4% B. 2.2% C. 2.2% D. 1.2% E. 4.7% F. 2.1% France: A. 3.2% B. 1.6% C. 2.2% D. 3.1% E. 5.9% F. 2.7% Germany: A. 0.2% B. 0.6% C.0.4% D. 0.6% E.3.8% F. 2.2% Italy: A. 0.8% B. 1.5% C. 1.2% D. 1% E. 5.1 % F. 2.1% Netherlands: A. 1.3% B. 1.7% C. 2.2% D. 3.1% E. 5% F. 1.6% Spain: A. 2% B. 1.7% C. 1% D. 0.9% E. 6.8 % F. 3.2% Ukraine: A. 6% B. 4% C. 3.2% D. 4.5% E. 19.2 % F. 10.6% Lebanon: A. 0.2% B. 0.4% C. 0% D. 0.7% E. 1.4 % F. 1.1% Nigeria: A. 0.2% B. 1.1% C. 0.2% D. 0% E. 2.2 % F. 0.6% Israel : A. 4.5% B. 2.9% C.- D. 1.1% E. 10.5 % F. 3.9% South Africa: A. 3.5% B. 8.5% C. 3.8% D. 0.5% E. 7.6 % F. 0%

Table B.2.1 cont. Overview of studies from which data was synthesised.

	Methodological summary						Statistical summary for cases with OA/joint pain	
Study ID [ID number]	**Country/ region**	**Design (response rate)**	**Setting**	**Mean age (SD): range**	**Grouping category (definition used by authors): description of cases ascertainment/time scale(anatomical site)**	**Sample (%F)**	**Mental health problem: assessment tool (reference standard), time frame**	**Prevalence rate of mental health problem in OA/ joint pain and compared groups**
(Continuous from previous page)								
Kadam, 2004 [32]	England and Wales/ national	Nested case-control	Primary Care	≥ 50	Medical records defined OA (OA):codes used in patients records (all)	Total: 23155 (65.5%) OA: 11375 Without OA: 11780	Depressive disorder: computerized patients records (code E2B)	In OA: 2.8% In without OA group: 1.9% OR 1.45(95% CI 1.15,1.82)
McWilliams, 2008 [33]	U.S.A/ national	Cross-sectional (BL: 81%)	General Population	≥18	Arthritis (arthritis): self-reported physician-diagnosed arthritis/ the last 12 months (all)	Total: 43093 Joint problems: 7876 (63.5%)	Panic with and without agoraphobia, social and simple phobia, GAD,MD, Dysthymia: AUDADIS (DSM-IV),12 months	In joint problems: Major depression: 9.8% Dysthymia: 3.0% Panic with agoraphobia: 0.8% Panic disorder: 2.6% Social phobia: 3.6% Simple phobia: 9.4% Generalized anxiety disorder: 3.5%
Moussavi, 2007 [34]	International/ 60 countries	Cross-sectional (BL: 63%-99%)	General Population	≥18	Arthritis (arthritis): self-reported, doctor-diagnosed or being treated for arthritis/ the last 12 months (all)	Total: 254404 Arthritis: 10431	Depressive episode: WHO-CIDI-III (DSM-IV), 12 months	In arthritis: 10.7% (95% CI 9.1- 12.3)
Fuller-Thompso, 2007 [35]	Canada/ national	Cross-sectional (BL: 83.7%)	General Population	≥20	Arthritis (arthritis): self-reported, health-professional diagnosed arthritis/rheumatism, excluding fibromyalgia/ lasted or are expected to last ≥6 months(all)	Total: 130880 Arthritis: 23405 (65.5%)	Major depression: WHO-CIDI-SF(DSM-III-R),12 months	In arthritis: Total: 9.9%, F: 10.8%, M: 8.2%
Patten, 2006 [36]	Canada/ national	Cross-sectional (BL: 77%)	General Population	≥ 15 (but group with arthritis was 25+)	Arthritis (arthritis): self-reported , doctor diagnosed presence of arthritis or rheumatism, excluding fibromyalgia/ ≥6 months (all)	Total: 36984 Joint problems: 8245	Panic disorder, social phobia, major depression: CCHS 1.2 interview based on WMH WHO-CIDI (DSM-IV), 12 months	In joint problems: Panic disorder: 3.0% (95% CI 2.5-3.6) Social phobia: 2.9% (95% CI 2.4-3.4) Major depression: 5 % (95% CI 4.3-5.7)
Wells, 1989a [37]	U.S.A/ LA ECA site	Cross-sectional (BL: 68%)	General Population	Total: 39.5	Arthritis (arthritis): indicated on a scale if is ever/ currently under a medical care for arthritis/current and lifetime (all)	Total: 2554 (50.4%) Lifetime (current) arthritis: 518 (417)	Anxiety Disorders: NIMH-DIS, recent and lifetime (DSM-III)	In lifetime arthritis: Lifetime 20 % (SE 2.5) Recent 9.1 % (SE 1.8) With current arthritis: Lifetime 20.7% (SE 3.3) Recent 11.9 % (SE 2.6)

B.3 QUALITY ASSESSMENT

Table B.3.1 Summary of quality assessment of 127 quality assessed articles.

Authors	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	Scores
Included studies:																
Allen et al. (2008)	+	+	-	-	+	+	+	+	+	+	+	+	+	+	+	13
Barberger-Gateau et al. (1992)	+	+	-	?	-	+	+	-	-	+	+	+	+	+	-	9
Creamer et al. (1999)	+	+	-	-	+	?	+	-	+	+	+	+	+	+	+	11
Croft, Jordan, Jinks (2005)	+	+	-	-	+	?	+	+	-	+	+	+	+	+	+	11
Dexter & Brandt (1994)	+	+	-	-	+	+	+	?	?	+	+	+	+	+	+	11
Dunlop et al. (2004)	+	+	-	-	-	?	-	-	-	+	?	+	+	+	-	6
Dunlop et al. (2005)	+	+	-	-	-	?	-	+	-	+	?	+	+	+	-	7
Figaro et al.(2005)	+	+	-	+	-	-	+	+	+	+	+	+	+	+	+	12
Fisher et al.(2004)	+	+	-	-	-	+	+	+	-	-	-	+	+	+	+	9
Fuller-Thompson & Shaked (2009)	+	+	-	-	-	?	-	+	-	+	+	+	+	+	+	10
Gureje et al. (2008)	+	+	-	-	-	+	+	+	+	+	+	+	+	?	-	10
Hill, Dziedzic, Thomas et al. (2007)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
He et al. (2008)	+	+	-	-	-	+	+	+	-	+	-	+	+	+	+	10
Jakobsson & Hallberg (2006)	+	+	-	-	-	-	-	-	+	+	+	-	+	+	+	8
Kadam et al. (2004)	+	+	-	+	+	+	+	N/A	N/A	+	+	-	+	+	+	13
Kramer et al. (2002)	+	+	-	?	-	+	+	-	+	+	+	+	+	+	?	10
Leveille et al. (2007)	+	+	-	+	+	+	+	+	-	+	+	+	+	+	+	13
Mallen & Peat (2008)	+	+	+	+	?	+	-	+	-	+	+	+	+	+	-	11
McWilliams et al. (2008)	+	+	-	?	?	+	+	+	-	+	+	+	+	?	-	9
Memel et al. (2000)	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	14
Moussavi et al. (2007)	+	+	-	-	?	+	+	?	-	+	?	+	+	+	-	8
Muus et al. (2007)	+	+	-	-	-	+	+	?	-	+	+	-	-	+	+	8
Niti et al.(2007)	+	+	-	+	-	+	+	+	+	+	?	+	+	+	+	12
Nour et al. (2005)	+	+	-	+	+	+	+	?	-	+	+	+	+	+	+	12
O'Reilly et al. (1998)	+	+	-	+	+	+	+	-	-	+	+	+	+	+	-	11
Patten et al. (2006)	-	?	-	+	-	+	+	+	-	+	+	+	+	?	-	8
Peat, Thomas, Handy et al. (2006a)	+	+	+	?	+	+	-	+	+	+	+	+	+	+	+	13
Polsky et al. (2005)	+	+	-	?	?	+	+	+	-	-	+	+	+	+	+	10
Rosemann, Backenstrass, Joest et al. (2007a)	+	-	-	+	+	+	?	+	+	+	+	+	+	+	+	12

Table B.3.1 Summary of quality assessment of 127 quality assessed articles.

Authors	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	Scores
Sale et al. (2008)	+	+	-	-	+	+	-	+	+	+	?	+	+	+	+	11
Schram et al. (2008)	-	?	-	-	-	+	-	+	+	+	+	+	+	?	+	8
Scudds & Robertson (2000)	+	+	-	-	+	+	+	+	-	+	+	+	+	+	+	12
Szoeke et al. (2008)	+	+	-	-	+	+	?	+	-	+	+	+	-	+	-	9
Wells et al. (1989a)	+	+	+	-	-	+	+	-	-	+	+	+	+	+	+	11
Wilcox et al. (2000)	+	+	-	+	+	+	+	-	?	+	+	+	+	+	+	12
Wilkie et al. (2007)	+	+	+	+	+	+	-	+	-	+	+	+	+	+	+	13
Woo et al. (1994)	+	+	-	-	-	+	+	-	-	+	+	+	+	+	+	10
Excluded multiple articles:																
Allen et al. (2009)	+	+	-	-	+	+	+	+	+	+	+	+	+	+	+	13
Elliot, Kraus, Fang et al. (2007)	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	14
Jinks, Jordan, Blagojevic, Croft (2008)	+	+	-	?	+	+	-	+	+	+	+	+	+	+	+	12
Jinks, Jordan, Croft (2002)	+	+	?	-	+	+	-	+	-	+	+	+	+	+	+	11
Jordan et al. (2006)	+	+	-	?	+	+	-	+	-	+	+	+	+	+	+	11
Lee et al. (2007)	+	+	?	-	-	+	+	+	-	+	+	+	+	+	+	11
Mallen et al. (2007)	+	+	+	+	+	+	-	-	-	+	+	+	+	+	+	12
Munce & Stewart (2007)	+	+	?	+	-	+	?	+	-	+	+	+	+	+	+	11
Peat & Thomas (2009)	+	+	+	-	+	+	-	+	+	+	+	+	+	+	-	12
Peat, Thomas, Croft (2006c)	+	+	+	+	+	+	-	-	+	+	+	+	+	+	-	12
Rosemann, Kuehlein, Laux, Szecsnyi (2007)	+	+	-	+	+	+	?	+	+	+	+	+	+	+	+	13
Rosemann, Wensing, Szecsnyi, Grol (2009)	+	+	-	+	+	+	?	+	+	+	+	+	+	+	+	13
Rosemann, Laux, Szecsnyi, Grol (2008)	+	+	-	+	+	+	?	+	+	+	+	+	+	+	+	13
Rosemann, Gensichen, Sauer et al. (2007)	+	+	-	-	+	+	?	+	+	+	+	+	+	+	+	12
Rosemann, Grol, Herman et al. (2008)	+	+	-	-	+	+	?	+	+	+	+	+	+	+	+	12
Rosemann, Joos, Koerner et al. (2006)	+	+	-	+	+	+	-	+	+	+	+	+	+	+	-	12
Rosemann, Kuehlein, Laux, Szecsnyi (2008)	?	+	-	+	+	+	?	+	+	+	+	+	+	+	+	12
Rosemann, Laux, Kuehlein (2007)	?	+	-	+	+	+	?	+	+	+	+	+	+	+	+	12
Rosemann, Laux, Szecsnyi et al. (2008)	?	+	-	+	+	+	?	+	+	+	+	+	+	+	+	12

Table B.3.1 cont. Summary of quality assessment of 127 quality assessed articles.

Authors	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	Scores
Rosemann, Joos, Szecsnyi et al. (2007)	+	+	-	+	+	+	?	+	+	+	+	+	+	+	+	13
Scott et al. (2009)	+	+	-	-	-	+	-	+	-	+	-	+	+	-	-	7
Sherman (2003)	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	14
Szoeke et al.(2005)	+	+	-	-	-	+	-	+	-	+	?	+	+	+	+	9
Thomas, Dunn, Mallen, Peat (2008)	+	+	+	+	+	+	-	-	-	+	+	+	+	+	+	12
Thomas, Peat, Mallen et al. (2008)	+	+	+	+	+	+	-	-	-	+	+	+	+	+	+	12
Verweij et al. (2009)	+	+	-	+	+	+	+	-	+	+	+	+	+	+	+	13
Wells et al. (1988)	+	+	+	-	-	+	+	-	-	+	+	+	+	+	+	11
Wood et al. (2007)	+	+	+	+	+	+	-	-	+	+	+	+	+	+	+	13
Mixed anxiety and depression:																
Alonso et al. (2004)	+	+	+	-	-	+	-	-	-	+	-	+	+	+	+	9
Antonopoulou et al. (2009)	+	+	+	+	?	+	+	+	-	+	?	+	+	+	-	11
Bot, Van Der Waal, Van Der Windt et al. (2005)	+	+	-	+	**+**	+	+	?	-	+	+	?	+	+	+	11
Bot, Van Der Waal, Terwee et al. (2005)	+	+	-	+	**+**	+	+	+	-	+	+	+	+	+	+	13
Boutron et al. (2008)	+	+	-	+	+	+	+	+	-	+	+	+	+	+	+	13
Busija et al. (2007)	+	+	-	?	-	+	+	-	+	+	+	+	+	+	+	11
Cimmino et al. (2005)	+	+	-	-	+	+	+	+	+	+	+	-	-	+	+	11
Fernandez-Lopez et al. (2008)	+	+	+	-	+	+	+	+	+	+	+	+	+	+	-	13
Hill, Parsons, Taylor, Leach (1999)	+	+	-	+	-	+	+	+	?	+	+	+	-	+	+	11
Hill, Gill, Taylor et al. (2007)	?	+	?	-	-	+	+	-	-	+	+	+	+	+	+	9
Hughes et al. (1994)	?	+	-	-	+	?	+	?	+	+	?	?	+	+	+	8
Jinks, Jordan, Croft (2007)	+	+	-	?	+	+	-	+	+	+	+	+	+	+	-	11
Jones et al. (2008)	+	+	-	+	?	?	?	?	-	+	+	+	+	-	+	8
Jordan et al. (2008)	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	14
Kaplan et al. (2003)	?	+	-	-	-	+	+	+	-	+	?	?	+	+	+	8
Lima et al. (2009)	+	+	+	-	-	+	+	?	-	+	+	+	+	+	+	11
Machado et al. (2006)	+	+	-	-	?	+	+	?	-	-	-	+	+	+	+	8
Mitchell et al. (2006)	?	+	-	+	+	+	+	-	-	+	+	+	+	+	-	10

Table B.3.1 cont. Summary of quality assessment of 127 quality assessed articles.

Authors	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	Scores
Palmer et al. (2006)	+	+	-	?	+	+	+	-	+	+	+	+	+	+	-	11
Peat, Thomas, Handy, Croft (2004)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
Rannou et al. (2007)	+	+	-	+	+	+	-	?	-	-	+	+	+	+	-	9
Strine et al. (2004)	+	+	-	+	?	+	+	?	-	+	+	-	+	+	+	10
Tangtrakulwanich et al.(2006)	+	+	-	-	+	+	+	?	-	+	+	+	+	+	+	11
Van der Waal, Terwee, van der Windt et al. (2005)	+	+	-	+	+	+	-	+	-	+	+	+	+	+	+	12
Van der Waal, Bot, Terwee et al. (2005)	+	+	-	+	+	+	-	?	+	+	+	+	+	+	+	12
Van der Waal, Bot, Terwee et al. (2006)	+	+	+	+	+	+	-	?	-	+	+	+	+	+	+	12
Van der Windt, Kuijpers, Jellema et al . (2007)	+	+	-	+	?	+	-	?	-	+	+	+	+	+	+	10
Wang et al. (2008)	+	+	-	-	+	+	-	+	?	-	+	+	+	+	+	10
Central tendency only:																
Appelt et al .(2007)	+	-	-	-	+	-	+	+	-	+	+	+	+	+	+	10
Badcock et al. (2002)	+	+	-	-	+	+	+	+	+	+	+	+	+	?	-	11
Baker (2003)	+	+	-	-	?	+	+	+	-	+	?	+	+	+	+	10
Brandt, Heilman et al. (2000)	+	-	-	+	+	+	+	-	-	+	+	+	+	+	+	11
Ferreira & Sherman (2006)	?	?	-	-	-	?	-	?	-	+	+	+	-	+	+	7
Gignac et al. (2008)	+	+	-	+	-	?	+	+	+	+	+	-	+	+	+	11
Hampson et al. (1996)	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	14
Hopman-Rock et al. (1997)	+	+	-	?	?	+	+	+	?	+	+	?	+	-	+	9
Kalichman et al. (2007)	+	+	-	-	+	?	+	?	-	+	+	+	+	+	-	9
Maly, Costigan, Olney (2007)	+	+	-	+	+	-	-	?	-	+	?	+	+	+	-	8
Martin (1996)	+	+	-	+	+	-	-	?	-	+	?	+	+	+	+	9
Menz et al. (2006)	+	+	-	+	+	+	+	-	-	+	?	+	-	+	-	9
Reilingh et al. (2008)	+	+	-	+	+	+	-	?	-	+	+	+	+	+	+	11
Spies-Dorgelo et al. (2007)	+	+	-	+	+	+	-	+	+	+	+	+	-	+	+	12
Tsai (2005)	+	+	-	?	-	+	?	?	-	+	?	+	+	+	+	8
Viinamaki et al. (2002)	+	+	-	?	?	+	+	?	-	+	+	+	+	+	+	10
Williams et al. (2004)	+	+	-	+	+	+	?	?	-	+	+	+	+	+	-	10

Table B.3.1 cont. Summary of quality assessment of 127 quality assessed articles.

Authors	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	Scores
Excluded multiple articles with central tendency only:																
Ang et al. (2002)	+	+	-	-	+	-	-	+	-	+	?	+	+	+	+	9
Kuijpers et al. (2006)	+	+	-	+	+	+	-	?	-	+	+	+	-	+	+	10
Ferreira & Sherman (2007)	+	+	-	-	-	?	-	+	+	+	+	+	-	+	+	9
Maly, Costigan, Olney (2006)	+	+	-	?	+	+	-	?	-	+	+	+	+	+	+	10
Maly, Costigan, Olney (2006)	+	+	-	+	+	-	-	?	-	+	-	+	+	+	+	9
Maly, Costigan, Olney (2005)	+	+	-	-	+	-	-	?	-	+	-	+	+	+	+	8
Definition of OA illegible:																
McWilliams et al. (2004)	+	+	-	-	-	?	+	-	-	+	-	+	+	+	-	7
Required details unobtainable:																
Blay et al. (2007)	?	+	-	-	?	?	+	-	-	+	?	+	-	+	+	6
Chang-Quan et al. (2008)	?	+	-	-	-	?	-	?	-	+	?	+	-	+	-	4
Chou & Chi (2002)	+	+	-	-	-	+	-	-	+	+	+	+	-	+	+	9
Davis et al. (1992)	+	+	-	-	+	+	-	+	?	+	+	?	-	?	-	7
Fautrel et al. (2002)	+	+	-	-	-	+	?	?	-	+	+	+	-	+	+	8
Hawker et al. (2008)	+	?	-	+	+	+	-	?	-	+	+	-	-	+	+	8
McCauley et al. (2008)	+	+	-	+	-	?	+	+	+	+	+	+	-	+	+	11
Ormel et al. (1997)	+	+	-	?	-	+	+	-	-	+	-	+	-	+	+	8
Ormel et al. (1998)	?	+	-	?	-	+	+	-	-	+	-	+	-	+	-	6
Wu et al. (2004)	+	?	-	-	-	?	-	?	-	+	+	+	-	?	+	5

Note: - no, +- yes, ?- unclear; Bibliography contains only included studies.

B.4 OVERALL META-ANALYSES (STATA OUTPUTS)

Table B.4.1 Meta-analysis of log-transformed prevalence rates of 'moderate or worse' depression symptoms.

Study	ES	[95% Conf.	Interval]	% Weight
1	0.550	0.437	0.693	10.23
2	0.235	0.189	0.291	10.27
3	0.136	0.077	0.240	8.97
4	0.079	0.062	0.099	10.24
5	0.156	0.120	0.203	10.15
6	0.091	0.054	0.152	9.19
7	0.582	0.391	0.866	9.69
8	0.075	0.066	0.086	10.43
9	0.239	0.207	0.276	10.41
10	0.105	0.092	0.120	10.43
D+L pooled ES	0.171	0.110	0.266	100.00

Heterogeneity chi-squared = 387.40 (d.f. = 9) p = 0.000
I-squared (variation in ES attributable to heterogeneity) = 97.7%
Estimate of between-study variance Tau-squared = 0.4770

Test of ES=1 : z= 7.87 p = 0.000

Table B.4.2 Meta-analysis of log-transformed prevalence rates of 'mild or worse' depression symptoms.

Study	ES	[95% Conf.	Interval]	% Weight
1	0.153	0.124	0.189	5.71
2	0.188	0.166	0.212	6.01
3	0.351	0.290	0.426	5.78
4	0.238	0.203	0.278	5.91
5	0.106	0.060	0.188	3.83
6	0.550	0.452	0.671	5.75
7	0.412	0.300	0.568	5.18
8	0.200	0.156	0.258	5.52
9	1.174	0.758	1.817	4.54
10	0.235	0.176	0.313	5.34
11	0.370	0.342	0.400	6.11
12	0.397	0.346	0.455	5.97
13	0.271	0.236	0.310	5.97
14	0.346	0.245	0.487	5.06
15	0.340	0.285	0.407	5.83
16	0.218	0.170	0.279	5.54
17	0.326	0.296	0.359	6.08
18	0.661	0.561	0.779	5.88
D+L pooled ES	0.312	0.260	0.374	100.00

Heterogeneity chi-squared = 344.89 (d.f. = 17) p = 0.000
I-squared (variation in ES attributable to heterogeneity) = 95.1%
Estimate of between-study variance Tau-squared = 0.1379

Test of ES=1 : z= 12.62 p = 0.000

Table B.4.3 Meta-analysis of log-transformed prevalent rates of 'moderate or worse' anxiety symptoms.

Study	ES	[95% Conf.	Interval]	% Weight
1	0.205	0.187	0.225	24.68
2	0.299	0.240	0.371	18.79
3	0.264	0.213	0.328	18.82
4	0.323	0.230	0.453	13.30
5	0.271	0.245	0.299	24.40
D+L pooled ES	0.262	0.220	0.312	100.00

Heterogeneity chi-squared = 23.91 (d.f. = 4) p = 0.000
I-squared (variation in ES attributable to heterogeneity) = 83.3%
Estimate of between-study variance Tau-squared = 0.0299

Test of ES=1 : z= 15.03 p = 0.000

Table B.4.4 Meta-analysis of log-transformed prevalent rates of 'mild or worse' anxiety symptoms.

Study	ES	[95% Conf.	Interval]	% Weight
1	1.000	0.844	1.185	14.25
2	0.808	0.677	0.966	13.42
3	0.852	0.636	1.140	6.41
4	0.639	0.507	0.806	9.21
5	0.818	0.762	0.878	29.50
6	0.852	0.784	0.925	27.21
D+L pooled ES	0.833	0.768	0.903	100.00

Heterogeneity chi-squared = 10.12 (d.f. = 5) p = 0.072
I-squared (variation in ES attributable to heterogeneity) = 50.6%
Estimate of between-study variance Tau-squared = 0.0045

Test of ES=1 : z= 4.43 p = 0.000

Table B.4.5 Meta-analysis of log-transformed prevalent rates of major depression.

Study	ES	[95% Conf.	Interval]	% Weight
1	0.041	0.016	0.106	1.99
2	0.049	0.027	0.091	3.35
3	0.110	0.105	0.115	6.39
4	0.053	0.048	0.058	6.27
5	0.103	0.062	0.169	4.02
6	0.063	0.042	0.094	4.63
7	0.040	0.017	0.091	2.39
8	0.117	0.088	0.156	5.35
9	0.054	0.036	0.080	4.66
10	0.022	0.007	0.077	1.37
11	0.014	0.002	0.129	0.50
12	0.114	0.074	0.174	4.45
13	0.053	0.024	0.114	2.61
14	0.066	0.053	0.082	5.79
15	0.086	0.071	0.104	5.91
16	0.022	0.012	0.042	3.34
17	0.060	0.027	0.133	2.55
18	0.082	0.058	0.116	4.97
19	0.073	0.053	0.100	5.18
20	0.238	0.189	0.298	5.70
21	0.122	0.110	0.135	6.26
22	0.103	0.087	0.121	6.00
23	0.109	0.101	0.117	6.33
D+L pooled ES	0.079	0.067	0.093	100.00

Heterogeneity chi-squared = 342.54 (d.f. = 22) p = 0.000
I-squared (variation in ES attributable to heterogeneity) = 93.6%
Estimate of between-study variance Tau-squared = 0.1079

Test of ES=1 : z= 30.49 p = 0.000

Table B.4.6 Meta-analysis of log-transformed prevalent rates of dysthymia.

Study	ES	[95% Conf.	Interval]	% Weight
1	0.011	0.002	0.066	2.34
2	0.021	0.009	0.053	5.42
3	0.017	0.005	0.052	4.30
4	0.028	0.016	0.050	7.35
5	0.022	0.008	0.067	4.54
6	0.041	0.026	0.064	8.15
7	0.021	0.012	0.039	7.21
8	0.001	0.000	0.309	0.29
9	0.011	0.001	0.134	1.38
10	0.022	0.009	0.054	5.53
11	0.016	0.004	0.063	3.52
12	0.025	0.018	0.034	8.83
13	0.006	0.002	0.019	4.17
14	0.003	0.000	0.086	0.81
15	0.000	0.000	0.999	0.12
16	0.033	0.021	0.052	8.18
17	0.119	0.089	0.159	9.05
18	0.037	0.029	0.049	9.17
19	0.031	0.027	0.035	9.64
D+L pooled ES	0.027	0.020	0.037	100.00

Heterogeneity chi-squared = 100.98 (d.f. = 18) p = 0.000
I-squared (variation in ES attributable to heterogeneity) = 82.2%
Estimate of between-study variance Tau-squared = 0.2653

Test of ES=1 : z= 22.38 p = 0.000

Table B.4.7 Meta-analysis of log-transformed prevalence rates of GAD.

Study	ES	[95% Conf.	Interval]	% Weight
1	0.028	0.009	0.087	3.33
2	0.013	0.004	0.042	3.36
3	0.010	0.002	0.043	2.34
4	0.033	0.020	0.057	7.48
5	0.002	0.000	0.071	0.47
6	0.046	0.030	0.071	8.52
7	0.008	0.003	0.021	4.08
8	0.026	0.008	0.083	3.34
9	0.002	0.000	0.824	0.16
10	0.013	0.004	0.041	3.37
11	0.015	0.004	0.061	2.48
12	0.040	0.031	0.052	10.07
13	0.002	0.000	0.015	1.39
14	0.036	0.013	0.099	4.03
15	0.037	0.022	0.060	7.81
16	0.020	0.011	0.035	7.11
17	0.064	0.044	0.094	9.04
18	0.063	0.051	0.077	10.53
19	0.036	0.032	0.041	11.07
D+L pooled ES	0.031	0.024	0.039	100.00

Heterogeneity chi-squared = 69.91 (d.f. = 18) p = 0.000
I-squared (variation in ES attributable to heterogeneity) = 74.3%
Estimate of between-study variance Tau-squared = 0.1422

Test of ES=1 : z= 27.40 p = 0.000

Table B.4.8 Meta-analysis of log-transformed prevalence rates of social phobia.

Study	ES	[95% Conf.	Interval]	% Weight
1	0.000	0.000	1.2e+04	0.02
2	0.022	0.009	0.055	5.03
3	0.030	0.026	0.034	10.60
4	0.038	0.018	0.083	5.83
5	0.022	0.012	0.043	6.75
6	0.004	0.000	0.050	1.04
7	0.012	0.005	0.027	5.61
8	0.000	0.000	7409.655	0.02
9	0.000	0.000	1.9e+07	0.01
10	0.056	0.031	0.100	7.28
11	0.022	0.007	0.071	3.67
12	0.049	0.039	0.063	10.00
13	0.002	0.000	0.015	1.54
14	0.000	0.000	9390.202	0.02
15	0.040	0.024	0.064	8.09
16	0.010	0.005	0.022	5.64
17	0.033	0.020	0.055	7.86
18	0.081	0.067	0.098	10.33
19	0.037	0.033	0.042	10.64
D+L pooled ES	0.031	0.024	0.041	100.00

Heterogeneity chi-squared = 118.84 (d.f. = 18) p = 0.000
I-squared (variation in ES attributable to heterogeneity) = 84.9%
Estimate of between-study variance Tau-squared = 0.1767

Test of ES=1 : z= 25.01 p = 0.000

Table B.4.9 Meta-analysis of log-transformed prevalence rates of panic with agoraphobia.

Study	ES	[95% Conf.	Interval]	% Weight
1	0.003	0.000	0.090	1.46
2	0.022	0.009	0.055	5.88
3	0.033	0.015	0.075	6.07
4	0.016	0.008	0.034	6.27
5	0.006	0.001	0.048	2.98
6	0.030	0.018	0.050	6.86
7	0.015	0.007	0.031	6.37
8	0.009	0.001	0.062	3.25
9	0.004	0.000	0.245	1.07
10	0.030	0.014	0.065	6.21
11	0.017	0.005	0.064	4.66
12	0.027	0.019	0.037	7.27
13	0.011	0.005	0.026	5.93
14	0.000	0.000	9390.202	0.06
15	0.093	0.067	0.129	7.25
16	0.017	0.009	0.032	6.65
17	0.042	0.026	0.066	7.01
18	0.053	0.042	0.066	7.40
19	0.008	0.006	0.010	7.37
D+L pooled ES	0.022	0.014	0.035	100.00

Heterogeneity chi-squared = 199.35 (d.f. = 18) p = 0.000
I-squared (variation in ES attributable to heterogeneity) = 91.0%
Estimate of between-study variance Tau-squared = 0.7258

Test of ES=1 : z= 16.36 p = 0.000

Table B.4.10 Meta-analysis of log-transformed prevalence rates of PTSD.

Study	ES	[95% Conf.	Interval]	% Weight
1	0.000	0.000	1.2e+04	0.05
2	0.013	0.004	0.042	6.31
3	0.005	0.001	0.039	3.25
4	0.031	0.018	0.054	9.88
5	0.006	0.001	0.048	3.03
6	0.010	0.004	0.025	7.79
7	0.010	0.004	0.024	7.79
8	0.001	0.000	0.309	0.51
9	0.007	0.000	0.159	1.55
10	0.003	0.000	0.032	2.52
11	0.031	0.011	0.083	7.11
12	0.042	0.032	0.054	11.42
13	0.000	0.000	0.853	0.21
14	0.008	0.001	0.063	3.01
15	0.004	0.001	0.018	5.12
16	0.010	0.004	0.022	8.26
17	0.048	0.031	0.074	10.62
18	0.055	0.044	0.069	11.57
D+L pooled ES	0.018	0.012	0.027	100.00

Heterogeneity chi-squared = 72.25 (d.f. = 17) p = 0.000
I-squared (variation in ES attributable to heterogeneity) = 76.5%
Estimate of between-study variance Tau-squared = 0.3773

Test of ES=1 : z= 18.88 p = 0.000

Table B.4.11 Meta-analysis of log-transformed prevalence rates of panic disorder.

Study	ES	[95% Conf.	Interval]	% Weight
1	0.031	0.030	0.032	49.98
2	0.027	0.026	0.027	50.02
D+L pooled ES	0.029	0.025	0.033	100.00

Heterogeneity chi-squared = 82.60 (d.f. = 1) p = 0.000
I-squared (variation in ES attributable to heterogeneity) = 98.8%
Estimate of between-study variance Tau-squared = 0.0107

Test of ES=1 : z= 48.22 p = 0.000

B.5 PUBLICATION BIAS ANALYSES (STATA OUTPUTS)

Table B.5.1 Results of publication bias analyses of questionnaire assessed prevalence rates.

STATA output

Number of studies = 10 Root MSE = 6.632

Std_Eff	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
slope	-2.350264	.5210055	-4.51	0.002	-3.551705	-1.148823
bias	4.563562	5.075614	0.90	0.395	-7.140824	16.26795

Construct: 'Moderate or worse' depression symptoms

Number of studies = 18 Root MSE = 4.625

Std_Eff	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
slope	-1.077688	.2026203	-5.32	0.000	-1.507224	-.6481519
bias	-.8662851	2.470181	-0.35	0.730	-6.102834	4.370264

Construct: 'Mild or worse depression' symptoms

Number of studies = 5 Root MSE = 2.25

Std_Eff	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
slope	-1.610284	.164093	-9.81	0.002	-2.132501	-1.088067
bias	3.064652	2.334053	1.31	0.281	-4.363346	10.49265

Construct: 'Moderate or worse' anxiety symptoms

Number of studies = 6 Root MSE = 1.583

Std_Eff	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
slope	-.1676395	.0816767	-2.05	0.109	-.3944103	.0591313
bias	-.2693585	1.374855	-0.20	0.854	-4.086568	3.547851

Construct: 'Mild or worse anxiety' symptoms

Table B.5.2 Results of publication bias analyses of clinical interview assessed prevalence rates.

STATA output	Construct

Number of studies = 23 Root MSE = 3.73

Std_Eff	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
slope	-2.201553	.0769298	-28.62	0.000	-2.361537	-2.041568
bias	-1.979141	1.040259	-1.90	0.071	-4.142478	.1841952

Major depression

Number of studies = 19 Root MSE = 2.292

Std_Eff	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
slope	-3.207467	.1518885	-21.12	0.000	-3.527924	-2.88701
bias	-1.096108	.7364898	-1.49	0.155	-2.649966	.4577493

Dysthymia

Number of studies = 19 Root MSE = 1.707

Std_Eff	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
slope	-3.038313	.1037814	-29.28	0.000	-3.257273	-2.819354
bias	-1.441039	.5454236	-2.64	0.017	-2.591782	-.2902956

GAD

Number of studies = 19 Root MSE = 2.515

Std_Eff	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
slope	-3.145241	.119896	-26.23	0.000	-3.3982	-2.892283
bias	-1.017197	.7596155	-1.34	0.198	-2.619846	.5854513

Social phobia

Number of studies = 19 Root MSE = 3.354

Std_Eff	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
slope	-3.374512	.3203648	-10.53	0.000	-4.050423	-2.698601
bias	-1.067535	1.261041	-0.85	0.409	-3.7281	1.593029

Panic with agoraphobia

Number of studies = 18 Root MSE = 1.189

Std_Eff	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
slope	-2.786855	.1178428	-23.65	0.000	-3.036671	-2.53704
bias	-2.313868	.3904681	-5.93	0.000	-3.141623	-1.486112

PTSD

B.6 SENSITIVITY ANALYSES

Table B.6.1 Summary of results for the non-standardised (I) and the sequential algorithm method (II).

Construct/ method of omitting	Number of pooled estimates	Pooled prevalence (95%CI)	I^2 (95%CI)	Omitted studies (I^2 after omitting)
Major depression	23	7.3 (6.3,8.5)	93.6 (91.6,95.1)	
I. $I^2 \leq 75\%$	17	7.0 (6.0,8.0)	61.3 (34.4,77.2)	Patten,2006 (87.0) Ukraine*(82.4) Fuller-Thomson, 2007 (82.1) Nigeria* (78.3), McWilliams, 2008 (78.0), Dunlop,2004 (61.3), Israel* (57.1), U.S.A* (39.5)
$I^2 \leq 50\%$	15	6.5 (6.7,7.5)	39.5 (0.0,67.2)	
II. $I^2 \leq 75\%$	19	8.3 (7.5, 9.1)	73.9 (59.0,83.4)	Patten, 2006 (87.0), Ukraine* (82.4), Nigeria* (78.6), New Zealand* (73.9), Italy* (70.3), Dunlop,2004 (68.3), Fuller-Thomson, 2007 (65.4), McWilliams, 2008 (54.5), U.S.A* (49.8)
$I^2 \leq 50\%$	14	7.0 (6.0,8.3)	49.8 (7.2,72.9)	
Dysthymia	19	2.6 (2.0, 3.6)	82.2 (73.2,88.1)	
I. $I^2 \leq 75\%$	-	-	-	Nigeria* (81.6), McWilliams, 2008 (81.3), U.S.A*(82.4), Israel *(83.6) Ukraine* (0.0)
$I^2 \leq 50\%$	14	2.3 (2.0,2.9)	0.0 (0.0,29.3)	
II. $I^2 \leq 75\%$	-	-	-	Ukraine* (23.9)
$I^2 \leq 50\%$	18	2.7 (2.3, 3.2)	29.3 (0.0,60.0)	
'Moderate or worse' depression symptoms	10	14.6 (9.9, 21.0)	97.7 (96.8,98.3)	
I. $I^2 \leq 75\%$	4	16.6 (13.6,20.1)	72.5 (22.2,90.3)	Nour, 2005 (97.6), Creamer, 1999 (96.4), Peat, 2006a (95.0), Kramer, 2002(94.1), Memel, 2000 (95.0), Wilkie, 2007 (72.5), Rosemann, 2007(71.8), Croft, 2005 (0.0)
$I^2 \leq 50\%$	2	13.2 (10.7,16.2)	0.0 (0.0,0.0)	
II. $I^2 \leq 75\%$	4	16.6 (13.6,20.1)	72.5 (22.2,90.3)	Creamer, 1999 (96.8), Peat, 2006a (95.7), Wilkie, 2007 (94.7), Kramer, 2002 (89.0), Nour, 2005 (81.4), Memel, 2000 (72.5), Mallen, 2008 (44.2)
$I^2 \leq 50\%$	3	18.4 (15.8,21.3)	44.2 (0.0,83.3)	

Table B.6.1 cont. Summary of results for the non-standardised (I) and the sequential algorithm method (II).

Construct/ method of omitting	Number of pooled estimates	Pooled prevalence (95%CI)	I^2 (95%CI)	Omitted studies (I^2 after omitting)
Mild or worse depression symptoms	18	23.8 (20.6,27.2)	95.1 (93.4,96.3)	
I. I^2 ≤75%	9	25.1 (23.3,27.1)	71.9 (44.6,85.7)	Nour, 2005 (94.8), Woo, 1994 (93.4), Leveille, 2007 (93.5), Allen,2008 (92.4), Barberger-Gateau, 1992 (87.9), Mallen, 2008(84.8), Niti, 2007(82.9), Wilcox, 2000 (81.1), Kramer, 2002 (71.9), O'Reilley,1998 (67.6), Memel, 2000 (70.7), Rosemann, 2007 (68.4), Peat, 2006a (46.3)
I^2 ≤50%	5	24.2 (22.4,26.1)	46.3 (0.0,80.3)	
II. I^2 ≤75%	7	22.8 (20.7,25.1)	74.8 (51.2,87.0)	Woo, 1994 (93.9), Barberger-Gateau, 1992 (92.4), Allen, 2008 (90.4), Nour, 2005 (88.6), Mallen, 2008 (86.2), Peat, 2006a (84.3), Rosemann, 2007 (81.1), Niti, 2007(76.9), Leveille, 2007 (74.8), Kramer, 2002 (68.7), Wilcox, 2000 (57.9), Sale, 2008 (36.7)
I^2 ≤50%	6	24.9 (23.0, 26.8)	36.7 (0.0,74.8)	
GAD	19	3.0 (2.5,3.9)	74.3 (59.6,83.6)	
I. I^2 ≤75%	18	4.1 (3.2,4.6)	74.1 (58.9,83.7)	McWilliams, 2008 (74.1), Italy* (70.4), Ukraine* (69.9), New Zealand* (71.8), Nigeria* (68.0), U.S.A* (29.7)
I^2 ≤50%	13	2.6 (2.0, 3.5)	29.7 (0.0,63.6)	
II. I^2 ≤75%	18	2.8 (2.2,3.6)	63.7 (39.9,78.1)	U.S.A* (63.7), Ukraine* (56.8), Italy* (46.9)
I^2 ≤50%	16	2.9 (2.3,3.6)	46.9 (4.9,70.3)	
Social phobia	19	3.0 (2.3,3.9)	84.9 (77.6,89.7)	
I. I^2 ≤75%	-	-	-	Spain* (84.2), McWilliams, 2008 (85.0), Italy*(84.6), New Zealand* (85.3), Nigeria* (85.0), Mexico* (86.0), U.S.A* (0.0)
I^2 ≤50%	12	2.9 (2.6,3.3)	0.0 (0.0,46.3)	
II. I^2 ≤75%	18	3.0 (2.4,3.7)	65.7 (43.5,79.1)	U.S.A* (65.7), Spain (60.1), New Zealand* (52.2), Nigeria* (41.8)
I^2 ≤50%	15	3.1 (2.6,3.8)	41.8 (0.0,68.4)	
Panic with agoraphobia	19	2.2 (1.4,3.4)	91.0 (87.4,93.5)	
I. I^2 ≤75%	17	2.2 (1.7,3.0)	66.7 (44.7,80.0)	McWilliams,2008 (79.4), South Africa* (66.7), U.S.A* (23.1)
I^2 ≤50%	16	2.2 (1.8,2.7)	23.1 (0.0,57.7)	
II. I^2 ≤75%	17	2.2 (1.7,3.0)	66.7 (44.7,80.0)	McWilliams, 2008 (79.4), South Africa* (66.7), U.S.A* (23.1)
I^2 ≤50%	16	2.2 (1.8,2.7)	23.1 (0.0,57.7)	

Table B.6.1 cont. Summary of results for the non-standardised (I) and the sequential algorithm method (II).

Construct/ method of omitting	Number of pooled estimates	Pooled prevalence (95%CI)	I^2 (95%CI)	Omitted studies (I^2 after omitting)
PTSD	18	1.8 (1.2,2.6)	76.5 (63.0,85.0)	
I. I^2 ≤75%	16	1.3 (0.0,2.2)	63.4 (37.3, 78.6)	New Zealand* (77.6), U.S.A* (63.4), Ukraine* (32.4)
I^2 ≤50%	15	1.2 (0.0,1.8)	32.4 (0.0,63.6)	
II. I^2 ≤75%	17	1.5 (0.0,2.0)	70.2 (51.2,81.6)	U.S.A* (70.2), New Zealand* (63.4),Ukraine* (32.4)
I^2 ≤50%	15	1.2 (0.0,1.8)	32.4 (0.0, 63.6)	
'Moderate or worse' anxiety symptoms	5	20.8 (18.0,23.8)	83.3 (62.0,92.6)	
I. and II.				
I^2 ≤75%	-	-		
I^2 ≤50%	4	21.6 (20.3,23.1)	0.0 (0.0,71.8)	Peat, 2006a (0.0)
'Mild or worse' anxiety symptoms	6	45.4 (43.4,47.5)	50.6 (0.0,80.4)	
I. I^2 ≤75%	5	45.5 (43.2,47.9)	60.0 (0.0,85.0)	Mallen, 2008 (60.0), O'Reilley, 1998 (35.2)
I^2 ≤50%	4	46.1 (44.4, 47.9)	35.2 (0.0,77.4)	
II. I^2 ≤75%	-	-	-	O'Reilley, 1998 (17.8)
I^2 ≤50%	5	45.9 (44.4,47.3)	17.8 (0.0,82.9)	

Note: Omitted studies are reported in an order of dropping with I^2 after removing the study presented in brackets; *- study by He (2008).

Appendix C: Self-report measures for depression and anxiety symptom screening and assessment in osteoarthritis: a narrative review of selected measurement properties

This appendix supports selected analyses described in chapter three

C.1 METHODS OF QUANTIFYING CONCURRENT VALIDITY OF SELF-REPORT MEASURES

Quantifying concurrent validity is based on a fourfold (two-by-two) table summarising a binary outcome (see Table C.1.1). The table allows estimating numbers of: (a) Individuals with the disease detected by the test (true positives) (b) Individuals without the disease with a positive test (false positives) (c) Individuals with the disease with a negative test (false negatives) and (d) Individuals without the disease with a negative test (true negatives) (Porta, 2008).

Table C.1.1 2x2 diagnostic accuracy table

		Disease		
		Present	**Absent**	
Test	**Positive**	True positive (a)	False positives (b)	a+b
	Negative	False negatives (c)	True negatives (d)	c+d
		a+c	b+d	a+b+c+d

Source: Porta et al., 2008

Table C.1.2 Methods of quantifying concurrent validity of self-report measures

Concept	Definition	Equation	Interpretation	Comments
Accuracy	A measure of the ability to correctly classify cases and non-cases	$\frac{a+d}{a+b+c+d}$	Frequency	Affected by prevalence A diagnosis for rare conditions may result in high sensitivity and specificity, but low accuracy
Diagnostic odds ratio (DOR)	A global index of diagnostic accuracy, defined as the ratio of odds of positive test in subjects with the targeted disorder to the odds in subjects with the targeted disorder	(a/c)/(b/d)	Unclear	Rule-in and -out accuracy cannot be interpreted Likewise sensitivity and specificity, is unaffected by the prevalence, but is responsive to characteristic of the disease (e.g. severity and comorbidities)
Likelihood ratios:	"*The probability that a given test result would occur in a person with the target disorder divided by the probability that the same result would occur in a person without that disorder*." (Porta, p.145, 2008)		Classification of magnitude of change in disease likelihood: LR+ or LR- >10.0 or <0.10 - large change 5-10 or 0.10-0.2- moderate change 2-5 or 0.2-0.5 - small change 1-2 or 0.5-1.0 - negligible change (Ahrens & Pigeot, 2005)	Allows for predicting probability of abnormality from the test and is prevalence independent (Deeks & Altman, 2004)
Likelihood ratio for positive test results (LR+)	The likelihood of a positive test in patients with the targeted disorder versus without the targeted disorder	((sensitivity)/(1 -specificty))		Unaffected by prevalence, but affected by the spectrum of the disease, unless adapting for prior probabilities through Bayes theorem
Likelihood ratio for negative test results (LR-)	The likelihood of a negative test in patients with the targeted disorder versus without the targeted disorder	((1-sensitivity)/ (specificity))		

Table C.1.2 cont. Methods of quantifying concurrent validity of self-report measures.

Concept	Definition	Equation	Interpretation	Comments
Predictive values:				
Positive predictive value (PPV)	The probability of having the disease in a person who has a disease (a true positive). Also known as predictive value of a positive test.	$\frac{a}{a+b}$	Used in combination with sensitivity and specificity to estimate rule-in and out accuracies	Predict particular test result from normality and abnormality, but are prevalence dependent (Deeks & Altman, 2004)
Negative predictive value (NPV)	The probability of not having the disease in a person who does not has a disease (a true negative). Also known as predictive value of a negative test.	$\frac{d}{c+d}$		Increases with the increase of the prevalence of the disease
Sensitivity and specificity:				Describe how normality and abnormality predict particular test result (Deeks & Altman, 2004)
Sensitivity (Sn.)	*"The probability that a diseased person (case) in the population tested will be identified as diseased by the test (syn: true positive probability). Sensitivity is thus the probability of correctly diagnosing a case or the probability that any given case will be identified by the test (syn: true-positive rate)*." (Porta, p.227, 2008)	$\frac{a}{a+c}$	Exemplar interpretations: - An optimal balance between Sn. and Sp. - A two-stage diagnostic process-maximum Sn. plus sp. ≥75% for rule-out accuracy (Lowe et al., 2004a)	Independent of prevalence Likely to decrease with a cut-off point Likely decrease when subclinical symptoms are assessed Is likely to be higher in secondary care than primary care and community samples
Specificity (Sp.)	*"The probability that a person without the disease (non-case) will be correctly identified as non-diseased by the test. It is thus the probability of correctly identifying a non-diseased person with a test (syn: true-negative probability)*." (Porta, p.227, 2008)	$\frac{d}{b+d}$		Independent of prevalence Likely to increase with a cut-off point Is likely to be higher in community and primary care samples than secondary care patients (due to higher number of comorbidities)

Table C.1.2 cont. Methods of quantifying concurrent validity of self-report measures.

Concept	Definition	Equation	Interpretation	Comments
Receiver Operating Characteristic curve (ROC)	The area under ROC curve (AUC) is a global index for measuring the diagnostic performance of the test.	A relationship between sensitivity and 1-specificty (false positives probability) (Kumar & Indrayan, 2011)	Range from 0.5 (chance) to 1.0 (perfect diagnostic accuracy) Cut-off point with the highest AUC is considered optimal. Exemplar interpretation for a summary ROC (Jones & Athanasious, 2005): 0.97 - 1.0 - excellent 0.93 - 0.96 - very good 0.75 - 0.92- good 0.50 - 0.74- reasonable but with deficiencies	Inference on screening and case-finding abilities is disenabled More useful to compare discriminating ability between two tests with quantifiable statistical difference
The clinical utility index (UI)	"*A measure of the clinical value of a diagnostic test taking into accounts both the accuracy of the test and its occurrence*". (Mitchell, p. 413, 2009) Two aspects are measures rule-in accuracy (case-finding ability) and rule-out accuracy (screening ability).	Case-finding ability: UI+= Sn x PPV Screening ability: UI-= Sp x NPV (can be weighted)	Landis & Koch (1977, cited in Mitchell, 2009): 0.93 - 1.0- near perfect value 0.81 - 0.92- excellent value 0.61 - 0.80- good value 0.41 - 0.60- fair value 0.21 - 0.40- slight value 0.0 - 0.20- minimal value	As involves predictive values can be affected by prevalence As to, PPV increases and NPV decrease with prevalence and UI+ and UI- change accordingly When possible weighting should be executed
Youden's index	A global index of diagnostic performance.	((sensitivity + specificity) -1) (Kumar & Indrayan, 2011)	0 (poor) to 1 (perfect) No agreed cut-off point, the highest Youden's index suggests the best cut-off point.	Affected by the factors affecting sensitivity and specificity (a spectrum of disease) A balance between sensitivity and specificity is ignored

Note: a- true positives; b- false positives; c- false negatives; d- true negatives.

C.2 SEARCH STRATEGY

Box C.2.1 Terms used to search EMBASE, MEDLINE, PsycInfo, CINHAL, CSA ILLUMINA and ISI Web of Knowledge.

EMBASE and EBSCO Health databases (MEDLINE, PsycInfo, CINHAL)

Text-word searched in all text

Data synthesis/ recommendations:

Review* OR systematic and review OR summary OR recommend*

Depression/anxiety:

depress* OR Anxiet* OR depressive symptom* OR depressed person OR depress* and symptom* OR anxious person OR anxiety symptom* OR affect and symptom* OR psychol* and distress OR HAD* OR hospital anxiety depression scale OR BDI OR beck a depression inventory OR PHQ OR patient and health and questionnaire OR GAD OR generalized anxiety disorder scale

Primary care:

primary and care OR general and practic* OR family and practic* OR family and medicine OR general and practition* OR family and practition*

Psychometric/diagnostic utility:

clinical utility OR psychometric* OR diagnos* OR validity OR reliability OR accuracy OR sensitivity OR specificity OR predictive value OR ROC OR screening OR false positive OR false negative OR logistic regression OR likelihood ratio OR accuracy OR attitude* OR belief OR perception OR clinical utility OR barrier*

Musculoskeletal complaints:

Joint diseases OR arthrit* OR musculoskeletal OR hand OR knee OR ankle OR foot OR shoulder OR elbow OR wrist OR hip OR pain and psych* OR musculoskelet* OR spondylosis OR osteoarthos* OR joint and stiff* OR joint and pain OR joint and diseas* OR osteoarthrit*)

Older adults:

Old* OR elderly OR age* OR geriat* OR adult*

Thesaurus/Mesh-terms

Depression/anxiety:

Depression OR Depressive disorder OR depressive disorder, major OR mental disease OR mixed anxiety and depression OR major depression, OR mood disorder OR mental health OR beck anxiety inventory OR hospital anxiety and depression scale

Box C.2.1 cont. Terms used to search EMBASE, MEDLINE, PsycInfo, CINHAL, CSA ILLUMINA and ISI Web of Knowledge.

Musculoskeletal complaints:

arthritis OR cartilage OR elbow OR elbow joint OR hand OR hand joint OR hip OR hip joint OR knee OR knee joint OR arthralgia OR shoulder OR shoulder join OR shoulder pain OR wrist OR wrist joint OR ankle OR ankle joint OR foot OR spondylosis OR cartilage degeneration OR elbow disease OR hip pain OR hip osteoarthritis OR hip disease OR knee disease OR knee osteoarthritis OR knee pain OR shoulder disease OR wrist disease OR ankle pain OR foot disease OR foot pain OR joint OR joint degeneration OR joint stiffness OR finger joint OR interphalangeal joint, exp carpometacarpal joint OR carpal joint OR atlantoocipital joint OR atlantoaxial joint OR acromioclavicular joint OR toe joint OR radioulnar joint OR sacroiliac joint OR subtalar joint OR sternocostal joint OR sternoclavicular joint OR tarsal joint OR tarsometatarsal joint OR temporomandibular joint OR zygapophyseal joint OR metacarpophalangeal joint OR metatarsophalangeal joint OR proximal interhalangeal joint OR patellofemoral joint OR cartilage diseases OR osteoarthritis OR osteoarthritis spine OR osteoarthritis hip OR osteoarthritis knee OR musculoskeletal diseases OR joint diseases OR pain/PX

Primary care:

general practice OR primary health care OR general practitioner OR family practice OR primary health care OR physicians family

Diagnostic/psychometric utility:

Sensitivity and specificity OR diagnosis OR predictive value of tests OR ROC Curve OR mass screening OR reproducibility of results OR false positive reactions OR false negative reactions OR logistic models OR psychometrics OR questionnaires OR rating scale OR depression inventory OR self-rating depression scale

CSA ILLUMINA and ISI Web of Knowledge

(depress*) AND (primary and care OR general and practice OR general and popul*) AND (diagnos* OR psychometric OR accuracy) AND (musculosk* OR pain OR osteoarthr* OR arthritis*) AND (perception OR attitude OR belief OR clinical and utility)

C.3 COPIES OF THE FULL VERISONS OF THE REVIEWD DEPRESSION MEASURES

BDI-21

Roche

V 0477

Beck Depression Inventory

CRTN: _____ CRF number: _____

Page 14 patient inits: _____

Baseline

BDI®-II

Date:

Name: _____ Marital Status: _____ Age: _____ Sex: _____

Occupation: _____ Education: _____

Instructions: This questionnaire consists of 21 groups of statements. Please read each group of statements carefully, and then pick out the **one statement** in each group that best describes the way you have been feeling during the **past two weeks, including today**. Circle the number beside the statement you have picked. If several statements in the group seem to apply equally well, circle the highest number for that group. Be sure that you do not choose more than one statement for any group, including Item 16 (Changes in Sleeping Pattern) or Item 18 (Changes in Appetite).

1. Sadness

0 I do not feel sad.
1 I feel sad much of the time.
2 I am sad all the time.
3 I am so sad or unhappy that I can't stand it.

2. Pessimism

0 I am not discouraged about my future.
1 I feel more discouraged about my future than I used to be.
2 I do not expect things to work out for me.
3 I feel my future is hopeless and will only get worse.

3. Past Failure

0 I do not feel like a failure.
1 I have failed more than I should have.
2 As I look back, I see a lot of failures.
3 I feel I am a total failure as a person.

4. Loss of Pleasure

0 I get as much pleasure as I ever did from the things I enjoy.
1 I don't enjoy things as much as I used to.
2 I get very little pleasure from the things I used to enjoy.
3 I can't get any pleasure from the things I used to enjoy.

5. Guilty Feelings

0 I don't feel particularly guilty.
1 I feel guilty over many things I have done or should have done.
2 I feel quite guilty most of the time.
3 I feel guilty all of the time.

6. Punishment Feelings

0 I don't feel I am being punished.
1 I feel I may be punished.
2 I expect to be punished.
3 I feel I am being punished.

7. Self-Dislike

0 I feel the same about myself as ever.
1 I have lost confidence in myself.
2 I am disappointed in myself.
3 I dislike myself.

8. Self-Criticalness

0 I don't criticize or blame myself more than usual.
1 I am more critical of myself than I used to be.
2 I criticize myself for all of my faults.
3 I blame myself for everything bad that happens.

9. Suicidal Thoughts or Wishes

0 I don't have any thoughts of killing myself.
1 I have thoughts of killing myself, but I would not carry them out.
2 I would like to kill myself.
3 I would kill myself if I had the chance.

10. Crying

0 I don't cry anymore than I used to.
1 I cry more than I used to.
2 I cry over every little thing.
3 I feel like crying, but I can't.

THE PSYCHOLOGICAL CORPORATION®
Harcourt Brace & Company
SAN ANTONIO
Orlando • Boston • New York • Chicago • San Francisco • Atlanta • Dallas
San Diego • Philadelphia • Austin • Fort Worth • Toronto • London • Sydney

_____ Subtotal Page 1

Continued on Back

0154018392
NR15645

cont. BDI-21

Roche

V 0477

Beck Depression Inventory

CRTN: _____ CRF number: _____

Page 15 patient inits: _____

Baseline

11. Agitation

0 I am no more restless or wound up than usual.
1 I feel more restless or wound up than usual.
2 I am so restless or agitated that it's hard to stay still.
3 I am so restless or agitated that I have to keep moving or doing something.

12. Loss of Interest

0 I have not lost interest in other people or activities.
1 I am less interested in other people or things than before.
2 I have lost most of my interest in other people or things.
3 It's hard to get interested in anything.

13. Indecisiveness

0 I make decisions about as well as ever.
1 I find it more difficult to make decisions than usual.
2 I have much greater difficulty in making decisions than I used to.
3 I have trouble making any decisions.

14. Worthlessness

0 I do not feel I am worthless.
1 I don't consider myself as worthwhile and useful as I used to.
2 I feel more worthless as compared to other people.
3 I feel utterly worthless.

15. Loss of Energy

0 I have as much energy as ever.
1 I have less energy than I used to have.
2 I don't have enough energy to do very much.
3 I don't have enough energy to do anything.

16. Changes in Sleeping Pattern

0 I have not experienced any change in my sleeping pattern.
1a I sleep somewhat more than usual.
1b I sleep somewhat less than usual.
2a I sleep a lot more than usual.
2b I sleep a lot less than usual.
3a I sleep most of the day.
3b I wake up 1–2 hours early and can't get back to sleep.

17. Irritability

0 I am no more irritable than usual.
1 I am more irritable than usual.
2 I am much more irritable than usual.
3 I am irritable all the time.

18. Changes in Appetite

0 I have not experienced any change in my appetite.
1a My appetite is somewhat less than usual.
1b My appetite is somewhat greater than usual.
2a My appetite is much less than before.
2b My appetite is much greater than usual.
3a I have no appetite at all.
3b I crave food all the time.

19. Concentration Difficulty

0 I can concentrate as well as ever.
1 I can't concentrate as well as usual.
2 It's hard to keep my mind on anything for very long.
3 I find I can't concentrate on anything.

20. Tiredness or Fatigue

0 I am no more tired or fatigued than usual.
1 I get more tired or fatigued more easily than usual.
2 I am too tired or fatigued to do a lot of the things I used to do.
3 I am too tired or fatigued to do most of the things I used to do.

21. Loss of Interest in Sex

0 I have not noticed any recent change in my interest in sex.
1 I am less interested in sex than I used to be.
2 I am much less interested in sex now.
3 I have lost interest in sex completely.

3 4 5 6 7 8 9 10 11 12 A B C D E

_____ Subtotal Page 2
_____ Subtotal Page 1
_____ Total Score

NR15645

Source: http://www.ibogaine.desk.nl/graphics/3639b1c_23.pdf

PHQ-9

Ealing NHS
Primary Care Trust

PHQ-9 Depression

Over the last 2 weeks, how often have you been bothered by any of the following problems? *(Use "✔" to indicate your answer)*	**Not at all**	**Several days**	**More than half the days**	**Nearly every day**
1. Little interest or pleasure in doing things.....	0	1	2	3
2. Feeling down, depressed, or hopeless.....	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much.....	0	1	2	3
4. Feeling tired or having little energy.....	0	1	2	3
5. Poor appetite or overeating.....	0	1	2	3
6. Feeling bad about yourself — or that you are a failure or have let yourself or your family down.....	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television.....	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed? Or the opposite — being so fidgety or restless that you have been moving .around a lot more than usual.....	0	1	2	3
9. Thoughts that you would be better off dead or of hurting yourself in some way.....	0	1	2	3
Column totals	___	+ ___	+ ____	+ ___

= *Total Score* _____

Source: www.iapt.nhs.uk/silo/files/phq9-and-gad7.doc

HADS

Name: Date:

Instructions: Doctors are aware that emotions play an important part in most illnesses. If your doctor knows about these feelings he or she will be able to help you more. This questionnaire is designed to help your doctor know how you feel. Read each item and **place a firm tick** in the box opposite the reply which comes closest to **how you have been feeling in the past week**. Don't take too long over your replies: your immediate reaction to each item will probably be more accurate than a long thought out response.

1. I feel tense or 'wound up':		**A**	**8. I feel as if I am slowed down:**	**D**	
Most of the time		3	Nearly all of the time	3	
A lot of the time		2	Very often	2	
Time to time, occasionally		1	Sometimes	1	
Not at all		0	Not at all	0	
2. I still enjoy the things I used to enjoy:	**D**		**9. I get a sort of frightened feeling like 'butterflies in the stomach':**		**A**
Definitely as much	0		Not at all		0
Not quite so much	1		Occasionally		1
Only a little	2		Quite often		2
Not at all	3		Very often		3
3. I get a sort of frightened feeling like something awful is about to happen:		**A**	**10. I have lost interest in my appearance:**	**D**	
Very definitely and quite badly		3	Definitely	3	
Yes, but not too badly		2	I don't take as much care as I should	2	
A little, but it doesn't worry me		1	I may not take quite as much care	1	
Not at all		0	I take just as much care as ever	0	
4. I can laugh and see the funny side of things:	**D**		**11. I feel restless as if I have to be on the move:**		**A**
As much as I always could	0		Very much indeed		3
Not quite so much now	1		Quite a lot		2
Definitely not so much now	2		Not very much		1
Not al all	3		Not at all		0
5. Worrying thoughts go through my mind:		**A**	**12. I look forward with enjoyment to things:**	**D**	
A great deal of the time		3	A much as I ever did	0	
A lot of the time		2	Rather less than I used to	1	
From time to time but not too often		1	Definitely less than I used to	3	
Only occasionally		0	Hardly at all	2	
6. I feel cheerful:	**D**		**13.I get sudden feelings of panic:**		**A**
Not at all	3		Very often indeed		3
Not often	2		Quite often		2
Sometimes	1		Not very often		1
Most of the time	0		Not at all		0
7. I can sit at ease and feel relaxed:		**A**	**14. I can enjoy a good book or radio or TV programme:**	**D**	
Definitely		0	Often	0	
Usually		1	Sometimes	1	
Not often		2	Not often	2	
Not at all		3	Very seldom	3	

Scores of 0-7 in respective subscales are considered normal, with 8-10 borderline and 11 or over indicating clinical 'caseness'.

Source: http://www.sadness101.com/HAD.html

Note: A- anxiety items; D- depression items.

GAD-7

Ealing NHS
Primary Care Trust

GAD-7 Anxiety

Over the last 2 weeks, how often have you been bothered by the following problems? *(Use "✓" to indicate your answer)*	Not at all	Several days	More than half the days	Nearly every day
1. Feeling nervous, anxious or on edge	0	1	2	3
2. Not being able to stop or control worrying	0	1	2	3
3. Worrying too much about different things	0	1	2	3
4. Trouble relaxing	0	1	2	3
5. Being so restless that it is hard to sit still	0	1	2	3
6. Becoming easily annoyed or irritable	0	1	2	3
7. Feeling afraid as if something awful might happen	0	1	2	3

Column totals: ___ + ___ + ___ + ___

= *Total Score* _____

If you checked off any problems, how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?

Not difficult at all	Somewhat difficult	Very difficult	Extremely difficult
☐	☐	☐	☐

Source: http://www.antiagingvancouver.com/pdfs/anxiety-depression-questionnaire.pdf

C.4 DATA USED FOR ESTIMATING LIKELIHOOD RATIOS

Table C.4.1 Overview of data extracted to calculate likelihood ratios for BDI-II, PHQ-9 and -2, HADS-D.

Study ID	Questionnaire	N	Diagnosis (%)	Cut of point	Sensitivity (95% CI)	Specificity (95% CI)
Arroll, 2003	PHQ-2	476	Major depression (6.1%)	Yes to 1 or 2	97.0 (83.0,99.0)	67.0 (62.0,72.0)
				Yes to 1st	86.0 (69.0,95.0)	72.0 (67.0,76.0)
				Yes to 2nd	83.0 (66.0,92.0)	79.0 (74.0,82.0)
Arroll, 2010	PHQ-2	2,642	Major depression (6.2%)	PHQ-2:		
	PHQ-9			1	96.0	60.0
				2	86.0	78.0
				3	61.0	92.0
				4	40.0	96.0
				PHQ-9:		
				8	82.0	85.0
				10	74.0	91.0
				12	61.0	94.0
				15	45.0	97.0
				Depression algorithm	45.0	97.0
Axford, 2010	HADS-D	54	Depression disorder (27.7%)	HADS-D:		
	HADS			8^	53.0 (27.0,79.0)	85.0 (69.0,94.0)
				HADS:		
				16^	55.0 (32.0, 76.0)	88.0 (71.0, 96.0)
Beck, 1997	BDI-7	56	Major depression (22%)	6A	83.0	95.0
Bunevicius, 2007	HADS-D	503	Major depressive episode (22%)	6R	80.0	69.0
Corapcioglu, 2004	PHQ-9	1387	(A) Major + minor depressive disorder (19.3%)	Algorithm		
				(A)	76.0	85.3
			(B) Major depression (6.6%)	(B)	71.4	91.9
El-Rufaie, 1995	HADS-D	217	Depression (24%)	7B	66.0	97.0

Table C.4.1 cont. Overview of data extracted to calculate likelihood ratios for BDI-II, PHQ-9 and -2, HADS-D.

Study ID	Questionnaire	N	Diagnosis (%)	Cut of point	Sensitivity (95% CI)	Specificity (95% CI)
Harter, 2001	HADS	206	Major depression, dysthymia (11%)	16M	78.3	70.6
Kroenke, 2003	PHQ-2	580	Major depression (7%) Any depressive disorder (18.3%)	**Major depression:** 1 2 3N 4 5 6 **Any depressive disorder :** 1 2 3 4 5 6	 97.6 92.7 82.9 73.2 53.7 26.8 90.6 82.1 62.3 50.9 31.1 12.3	 59.2 73.7 90.0 93.3 96.8 99.4 65.4 80.4 95.4 97.9 98.7 99.8
Lam, 1995	HADS-D	100	Depressive disorders (9%)	6*	78.0	91.0
Li, 2007	PHQ-2	8205	Major depression in past 12 months (3.7%)	Yes to 1 and 2 Yes to 1 Yes to 2 Yes to 1 or 2Y	77.0 (72.6,82.9) 92.0 (88.6,95.8) 85.0 (80.2,89.8) 100	86.0(85.4,87.2) 79.0 (77.5,79.6) 85.0 (83.7,85.5) 77.0 (75.8,78.0)

Table C.4.1 cont. Overview of data extracted to calculate likelihood ratios for BDI-II, PHQ-9 and -2, HADS-D.

Study ID	Questionnaire	N	Diagnosis (%)	Cut of point	Sensitivity (95% CI)	Specificity (95% CI)
Phelan, 2010	PHQ-9	71	Major depression (12%) Major and minor depression (25%)	**Major depression:**		
				8	88.0 (56.0,98.0)	75.0 (71.0,77.0)
				9B	88.0 (56.0,98.0)	80.0 (76.0,82.0)
				10^	63.0 (33.0,85.0)	82.0 (78.0,85.0)
				11	63.0 (33.0,85.0)	84.0 (80.0,87.0)
				12	63.0 (33.0,85.0)	84.0 (80.0,87.0)
				Major and minor depression:		
				6	77.0 (56.0,90.0)	69.0 (63.0,74.0)
				7	77.0 (56.0,90.0)	77.0 (70.0,81.0)
				8B	77.0 (57.0,89.0)	83.0 (76.0,87.0)
				9	71.0 (51.0,85.0)	87.0 (80.0,91.0)
				10	59.0 (40.0,74.0)	89.0 (82.0,93.0)
Phelan, 2010	PHQ-2	71	Major depression (12%) Major and minor depression (25%)	**Major depression:**		
				1	88.0 (55.0,98.0)	61.0 (56.0,62.0)
				2B	75.0 (43.0,93.0)	67.0 (63.0,70.0)
				3^	63.0 (33.0,85.0)	85.0 (81.0,88.0)
				4	38.0 (15.0,62.0)	93.0 (91.0,97.0)
				5	38.0 (16.0,48.0)	98.0 (96.0,100)
				Major and minor depression:		
				1	82.0 (62.0,94.0)	67.0 (61.0,71.0)
				2	71.0 (50.0,86.0)	73.0 (66.0,78.0)
				3	53.0 (35.0,67.0)	90.0 (85.0,95.0)
				4	35.0 (21.0,40.0)	98.0 (94.0,100)
				5	18.0 (7.0, 22.0)	98.0 (95.0,100)

Table C.4.1 cont. Overview of data extracted to calculate likelihood ratios for BDI-II, PHQ-9 and -2, HADS-D.

Study ID	Questionnaire	N	Diagnosis (%)	Cut of point	Sensitivity (95% CI)	Specificity (95% CI)
Poole, 2009b	BDI-II	36	Any depression (72%)	14.5	100	60.0
				16.0	100	70.0
				17.5	89.0	90.0
				18.5	96.0	80.0
				20.0	92.0	80.0
				22.0*	89.0	90.0
				23.5	85.0	100
				24.5	73.0	100
Terluin, 2009	HADS-D	295	Any depressive disorder (49%) Moderate or severe depressive disorder (26%)	**Any depressive disorder:**		
				8	93.0	39.0
				9	88.0	49.0
				11B	75.0	67.0
				14	40.0	88.0
				Moderate or severe depressive disorder:		
				8	96.0	30.0
				10	87.0	48.0
				11	81.0	56.0
				12B	64.0	65.0
				15	39.0	85.0
Watts, 2002	HADS-D	115	Sub-clinical psychiatric (33.9%)	8^	24.0	89.0
Whooley, 1997	BDI-II	536	Major depression (18.1%)	BDI-II: 10^	89.0 (81.0,95.0)	64.0 (59.0,68.0)
				PHQ-2:		
				Yes to 1 or 2:		
				All	96.0 (90.0,99.0)	57.0 (53.0,62.0)
				By age:		
				18 to 35	100 (59.0,100)	59.0 (43.0,74.0)
				35-64	95.0 (88.0,99.0)	52.0 (46.0,58.0)
				65+	100 (54.0,100)	69.0 (60.0,77.0)

Table C.4.1 cont. Overview of data extracted to calculate likelihood ratios for BDI-II, PHQ-9 and -2, HADS-D.

Study ID	Questionnaire	N	Diagnosis (%)	Cut of point	Sensitivity (95% CI)	Specificity (95% CI)
Wilkinson & Barczak, 1988	HADS-D	100	Depressive disorders (23%)	8*	90.0	86.0
Yeung, 2008	PHQ-9	184	Major depression (20%)	15^	81.0	40.00

Note: Methods of ascertainment of optimal cut-off point, if such is specified: ^ - reported by authors as previously recommended; * - Area Under the Curve; A - balance between sensitivity, specificity, negative and positive predictive values; B - balance between sensitivity and specificity; M - minimised false positive and false negatives; N - sensitivity, specificity, positive predictive value, positive likelihood ratio, Area Under the Curve; Y - the Youden index.

Table C.4.2 Overview of data extracted to calculate likelihood ratios for HADS-A, GAD-7 and -2.

Authors	Questionnaire	N	Diagnosis (%)	Cut of point	Sensitivity (95% CI)	Specificity (95% CI)
Axford, 2010	HADS-A HADS	54	Anxiety disorders (31.48%)	HADS-A: 8^	88.0 (64.0,99.0)	81.0 (65.0,92.0)
				HADS: 16^	55.0 (32.0,76.0)	88.0 (71.0,96.0)
Bunevicius 2007	HADS-A	503	Anxiety disorders (27%) Social phobia (4%) Panic disorder (3%) GAD (25%)	**Anxiety disorders:**		
				9*	77.0	75.0
				Social phobia:		
				9*	95.0	63.0
				Panic disorder:		
				11*	100	77.0
				GAD		
				9*	76.0	73.0
El-Rufaie & Absood,1995	HADS-A	217	Anxiety (18.4%)	9B	66.0	93.0
Garcia-Campayo, 2010	GAD-7	212	GAD (50%)	GAD-7:		
				8	93.4	85.8
				9	91.5	90.6
				10^	86.8	93.4
				11	79.2	95.3
				12	74.5	96.2
				13	67.9	97.2
				14	62.3	100
Harter, 2001	HADS	206	Anxiety disorders (15%)	17M	75.0	72.2
Lam, 1995	HADS-A	100	Anxiety (6%)	3*	67.0	83.0

Table C.4.2 cont. Overview of data extracted to calculate likelihood ratios for HADS-A, GAD-7 and -2.

Authors	Questionnaire	N	Diagnosis (%)	Cut of point	Sensitivity (95% CI)	Specificity (95% CI)
Spitzer, 2006	GAD-7 GAD-2	965	GAD (7.6%) Panic disorder (6.8%)	**GAD:**		
				GAD-7		
				5	97.0 (90.0,100)	57.0 (53.0,60.0)
				6	95.0 (87.0,98.0)	65.0 (61.0,67.0)
				7	95.0 (87.0,98.0)	70.0 (67.0,73.0)
				8	92.0 (83.0,97.0)	76.0 (73.0,79.0)
				9	90.0 (81.0,96.0)	79.0 (76.0,82.0)
				10^	89.0 (80.0,95.0)	82.0 (80.0,85.0)
				15	48.0	95.0
				GAD-2:		
				2	95.0 (87.0,98.0)	64.0 (61.0,67.0)
				3K	86.0 (76.0,93.0)	83.0 (80.0,85.0)
				Panic disorder:		
				GAD-7:	94.0 (85.0,98.0)	56.0 (53.0,59.0)
				5	88.0(78.0,95.0)	64.0 (60.0,67.0)
				6	83.0(72.0,91.0)	69.0 (66.0,72.0)
				7	82.0 (70.0,90.0)	75.0 (72.0,78.0)
				8	79.0 (67.0,88.0)	78.0 (75.0,80.0)
				9	74.0 (62.0,84.0)	81.0 (78.0,83.0)
				10		
				GAD-2:		
				2	91.0 (81.0,97.0)	63.0 (60.0,66.0)
				3	76.0 (64.0,85.0)	81.0 (79.0,84.0)

Table C.4.2 cont. Overview of data extracted to calculate likelihood ratios for HADS-A, GAD-7 and -2.

Authors	Questionnaire	N	Diagnosis (%)	Cut of point	Sensitivity (95% CI)	Specificity (95% CI)
Spitzer, 2006	GAD-7 GAD-2	965	Social phobia (6.2%) PTSD (8.6%)	**Social phobia:**		
				GAD-7:		
				5	88.0 (85.0,98.0)	55.0 (52.0,59.0)
				6	87.0 (75.0,94.0)	63.0 (60.0,66.0)
				7	85.0 (73.0,92.0)	69.0 (66.0,72.0)
				8	78.0 (66.0,88.0)	74 .0 (71.0,77.0)
				9	77.0 (64.0,87.0)	77.0 (74.0,80.0)
				10	72.0 (59.0,83.0)	80.0 (77.0,83.0)
				GAD-2:		
				2	85.0 (73.0,93.0)	62.0 (59.0,65.0)
				3	70.0 (57.0,81.0)	81.0 (78.0,83.0)
				PTSD:		
				GAD-7:		
				5	90.0 (82.0,96.0)	57.0 (53.0,60.0)
				6	86.0 (76.0,92.0)	64.0 (61.0,68.0)
				7	78.0 (68.0,87.0)	70.0 (66.0,73.0)
				8	76.0 (65.0,85.0)	75.0 (72.0,78.0)
				9	74.0 (63.0,83.0)	78.0 (75.0,81.0)
				10	66.0 (55.0,76.0)	81.0 (78.0,84.0)
				GAD-2:		
				2	86.0 (76.0,92.0)	64.0 (60.0,67.0)
				3	59.0 (48.0,70.0)	81.0 (78.0,84.0)

Table C.4.2 cont. Overview of data extracted to calculate likelihood ratios for HADS-A, GAD-7 and -2.

Authors	Questionnaire	N	Diagnosis (%)	Cut of point	Sensitivity (95% CI)	Specificity (95% CI)
Spitzer, 2006	GAD-7 GAD-2	965	Any anxiety disorder (19.5%)	**Any anxiety disorder:**		
				GAD-7:		
				5	90.0 (85.0,94.0)	63.0 (60.0,66.0)
				6	85.0 (79.0,90.0)	71.0 (68.0,74.0)
				7	80.0 (74.0,86.0)	76.0 (73.0,79.0)
				8	77.0 (70.0,82.0)	82.0 (80.0,85.0)
				9	73.0 (66.0,80.0)	85.0 (83.0,81.0)
				10	68.0 (60.0,74.0)	88.0 (85.0,90.0)
				GAD-2:		
				2	86.0 (80.0,90.0)	70.0 (67.0,74.0)
				3S	65.0 (57.0,71.0)	88.0 (85.0,90.0)
Terluin, 2009	HADS-A	295	Any anxiety disorders (34%) Panic disorder, agoraphobia and social phobia (19%)	**Any anxiety disorders:**		
				8	98.0	27.0
				10	91.0	43.0
				11	84.0	51.0
				13B	65.0	70.0
				16	32.0	89.0
				Panic disorder, agoraphobia and social phobia:		
				8	100	23.0
				11	86.0	45.0
				13B	68.0	65.0
				16	41.0	87.0
Watts, 2002	HADS-A	115	Sub-clinical psychiatric diagnosis (33.9%)	8^	57.0	71.0
Wetherell, 2007	HADS-A	37	GAD (47.1%)	8^	96.7	66.0

Note: Methods of ascertainment of optimal cut-off point, if such is specified: ^ - reported by authors as previously recommended; * - Area Under the Curve; B - balance between sensitivity and specificity; K - sensitivity, specificity, LR+; M - minimalised false positive and false negatives; S - optimal for screening.

Appendix D: The course of anxiety and depression symptoms in older patients presenting to general practice with musculoskeletal pain

Part 2: A latent class growth analysis

This appendix supports selected analyses described in chapter five

D.1 DISTRIBUTION OF SCORES FOR CONTINUOUS BASELINE COVARIATES ASSESSED FOR ASSOCCIATIONS WITH CLUSTER MEMBERSHIP

Figure D.1.1 Distribution of the number of pain sites variable in 443 consenters for follow-up.

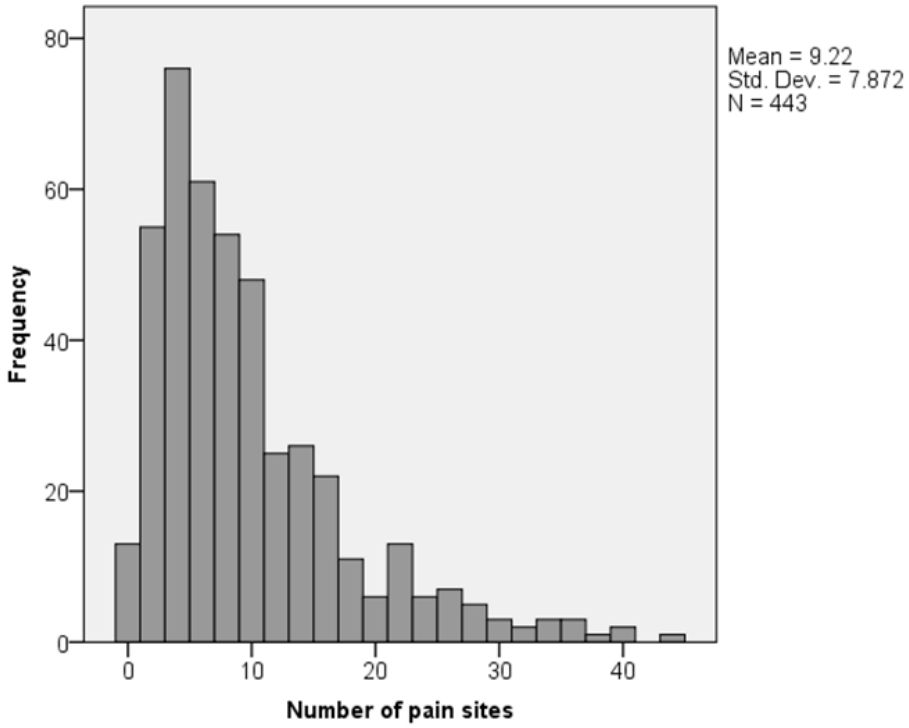


Figure D.1.2 Distribution of the pain interference with daily activities variable in 443 consenters for follow-up.

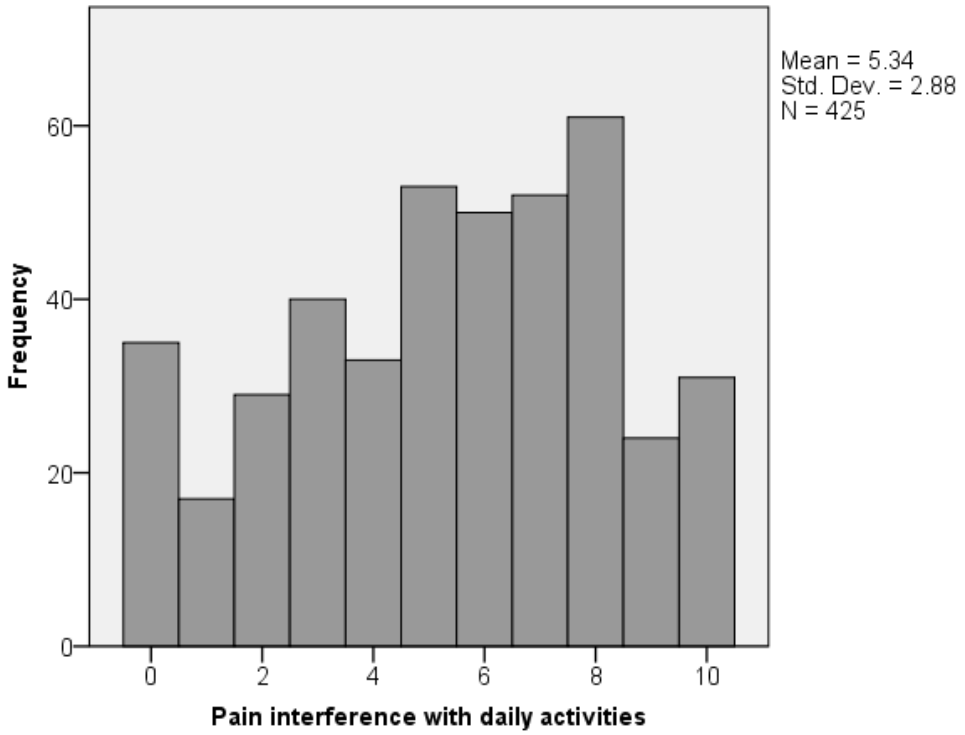


Figure D.1.3 Distribution of the pain interference with work variable in 443 consenters for follow-up.

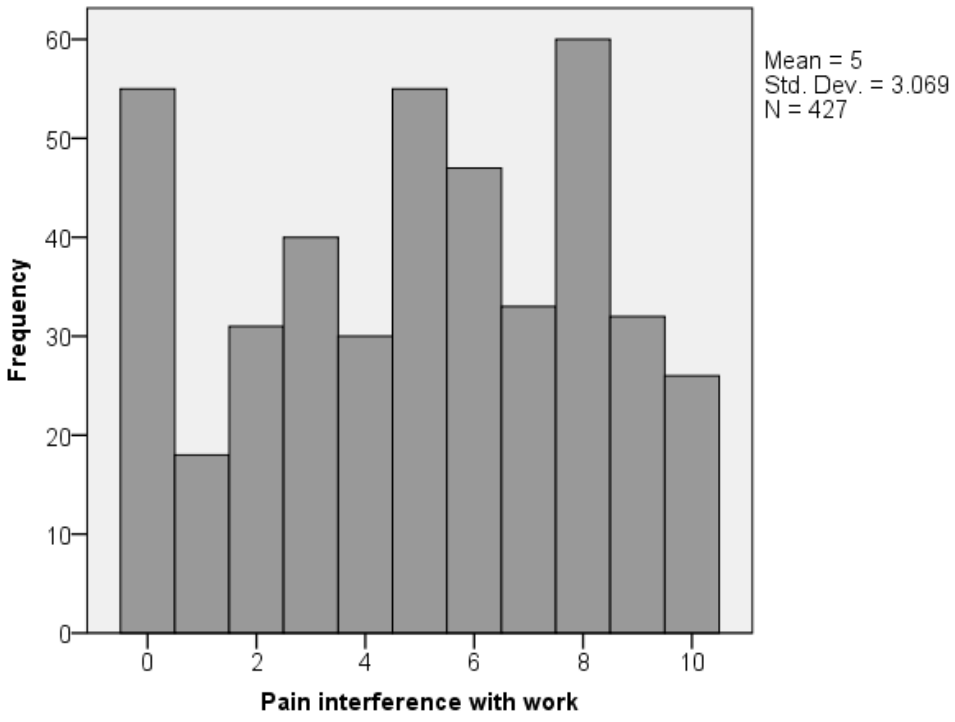
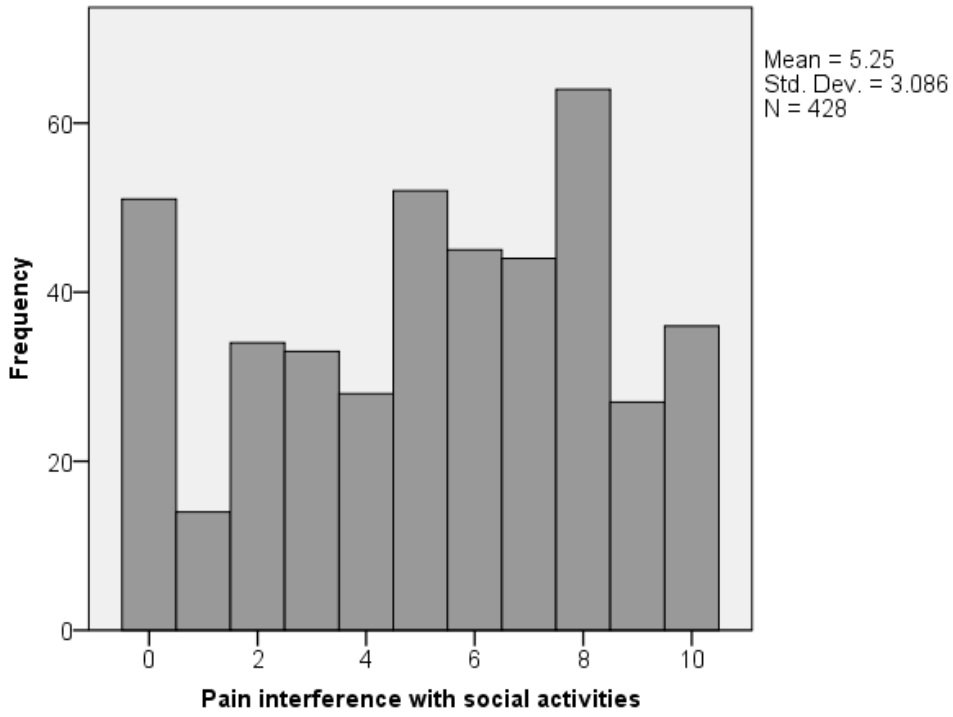
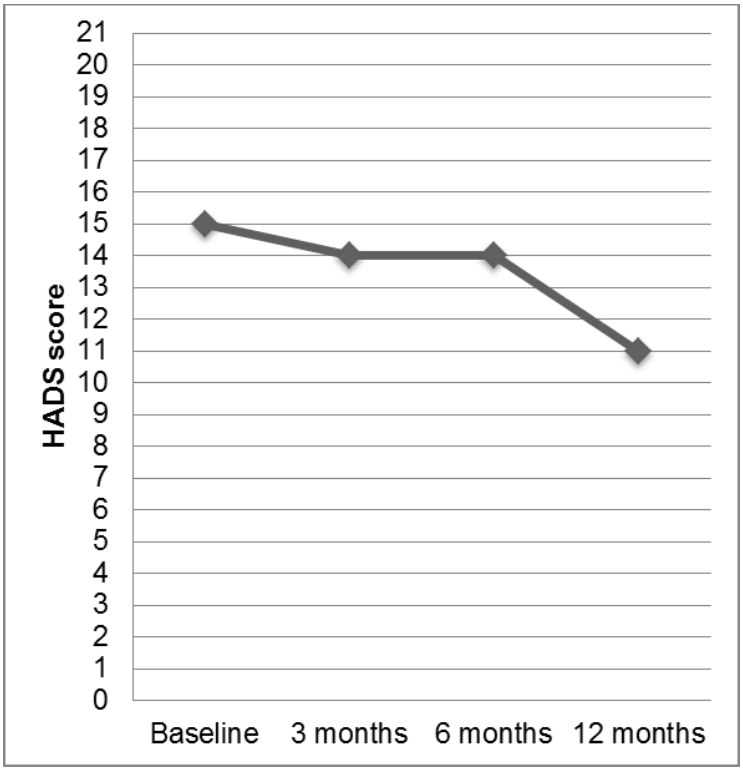


Figure D.1.4 Distribution of the interference with social activities variable in 443 consenters for follow-up.



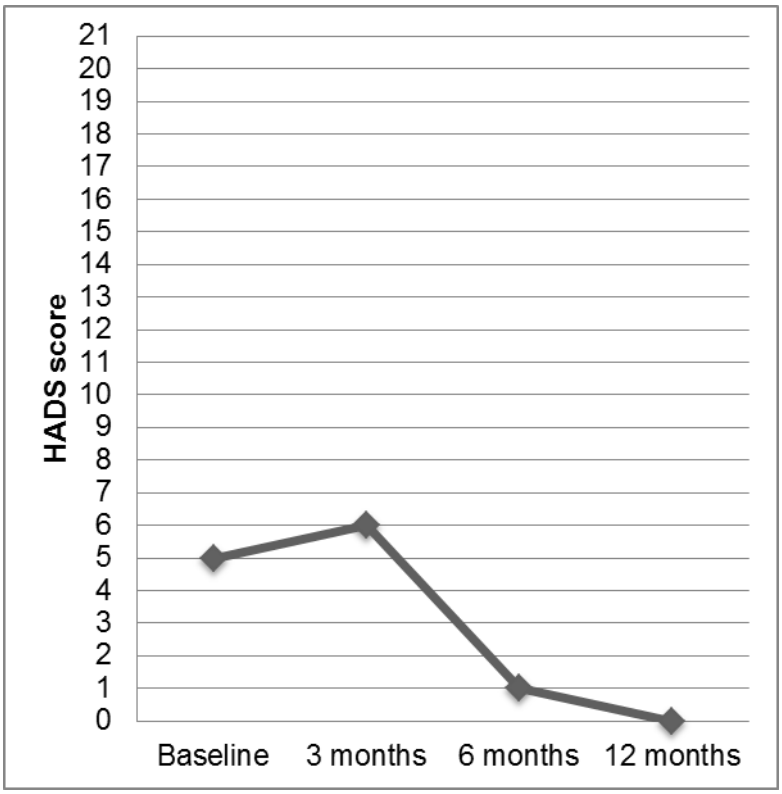
D.2 EXEMPLES OF PATTERNS OF HADS SCORES OVER TIME

Figure D.2.1 Exemplar for the pattern '1111'*.



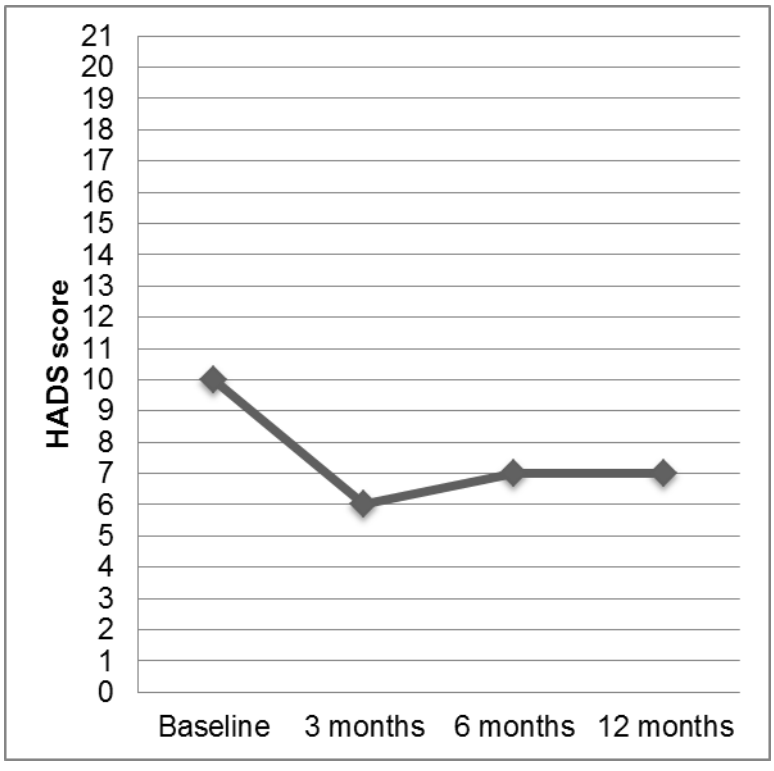
Note: *1- HADS score ≥ 8.

Figure D.2.2 Exemplar of the pattern '0000'*.



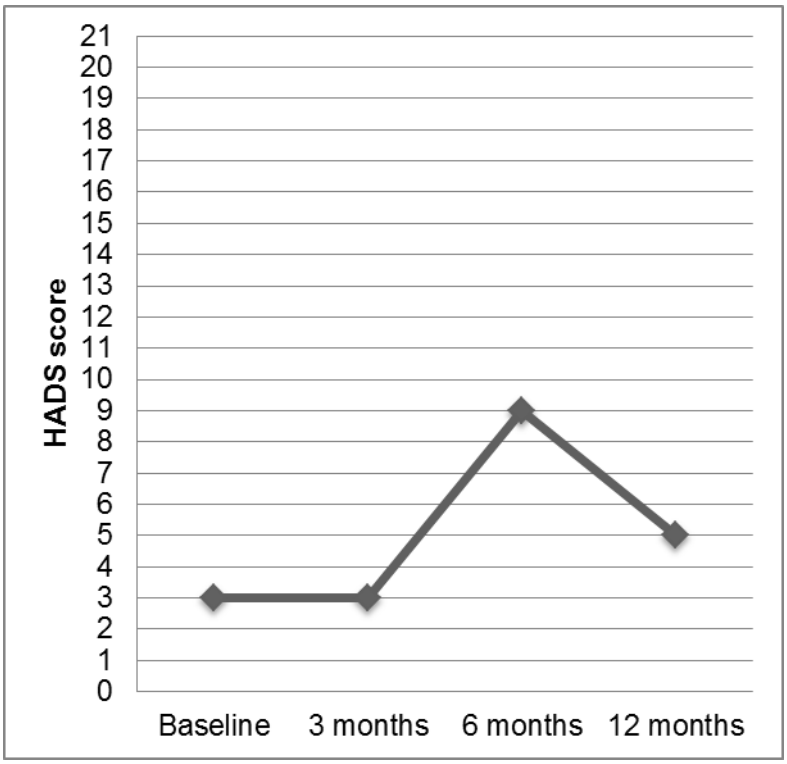
Note:*0- HADS score 0-7.

Figure D.2.3 Exemplar of the pattern '1000'*.



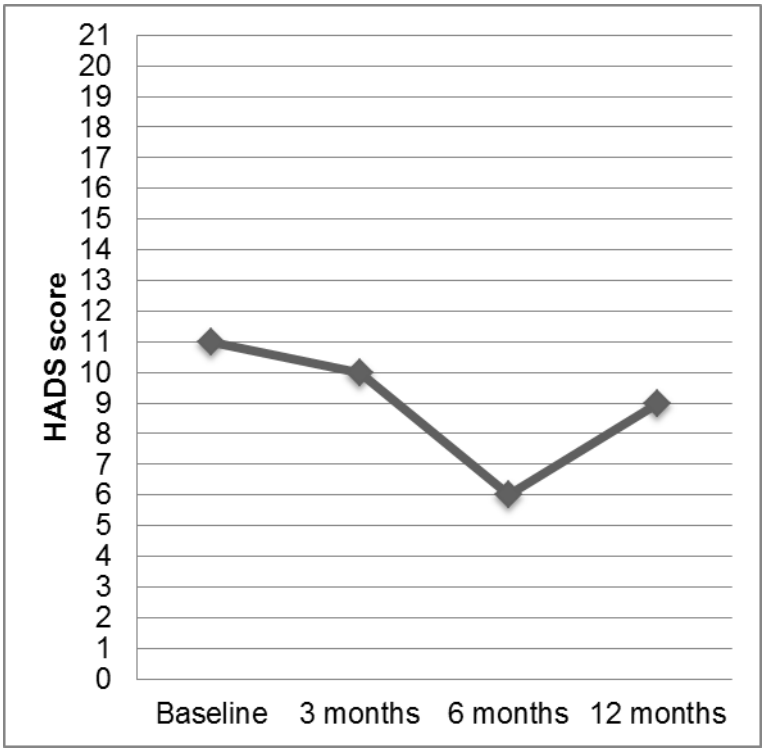
Note: *1- HADS score ≥ 8; 0- HADS score 0-7.

Figure D.3.4 Exemplar of the pattern '0010'*.



Note: *1- HADS score ≥ 8, 0- HADS score 0-7.

Figure D.3.5 Exemplar of the pattern '1101'*.



Note: *1- HADS score ≥ 8, 0- HADS score 0-7.

D.3 INDIVIDUAL DEPRESSION PATTERNS IN 2 CLUSTERS NESTED IN THE 2-CLUSTER LCGA DEPRESSION MODEL

Figure D.3.1 Individual HADS-D scores over time for 232 individuals in the *no depression symptom trajectory*.

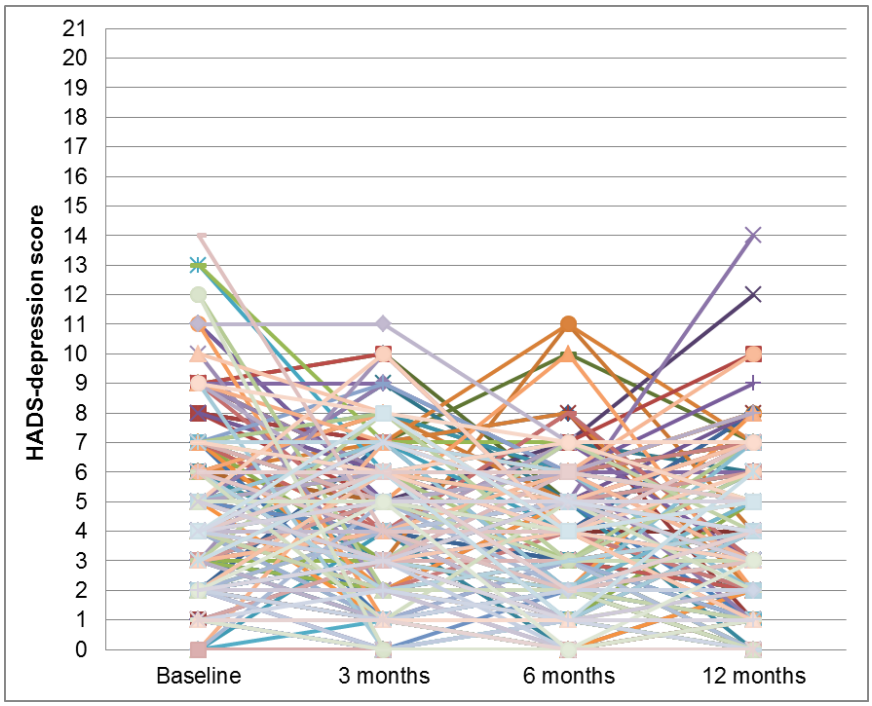
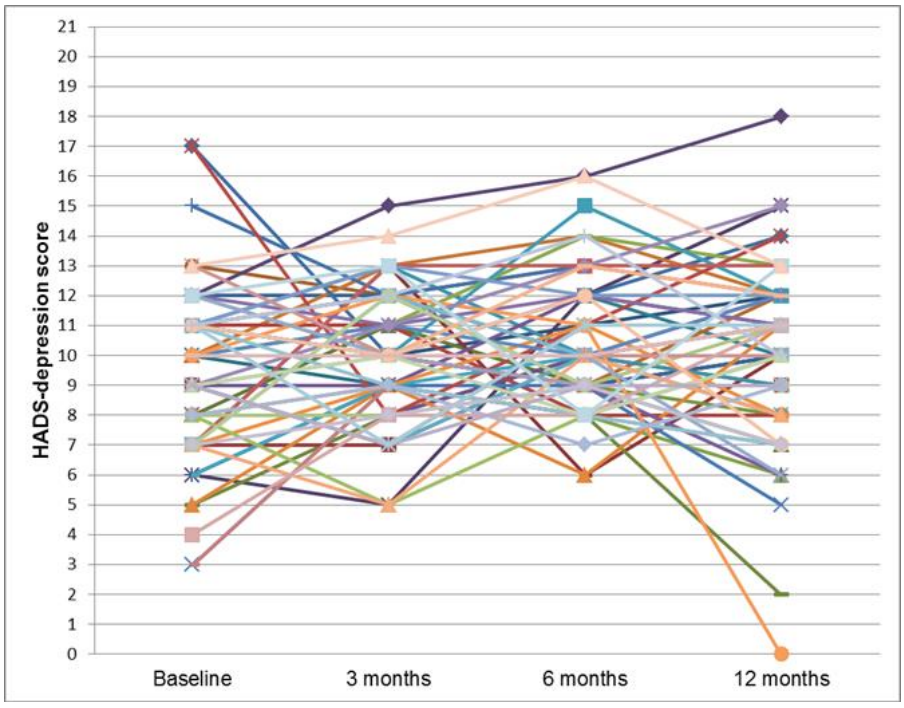


Figure D.3.2 Individual HADS-D score for 66 individuals in the *persistent depression symptom trajectory*.



D.4 THE 4-CLUSTER LCGA ANXIETY MODEL FOR 293 PARTICIPANTS WITH COMPLETE HADS-A DATA

Figure D.4.1 Cluster 1: 4-cluster LCGA anxiety model for complete binary anxiety data (n=121).

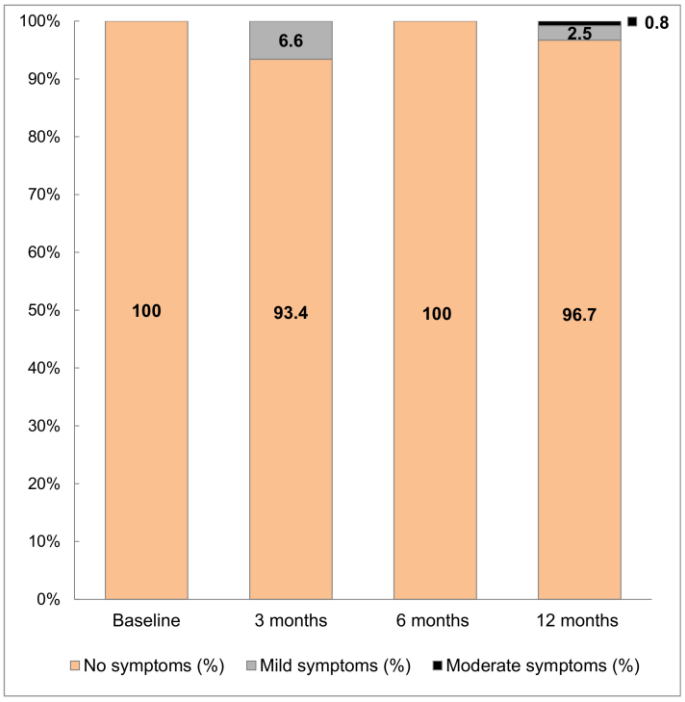


Figure D.4.2 Cluster 2: 3-cluster LCGA anxiety model for complete binary anxiety data (n=87).

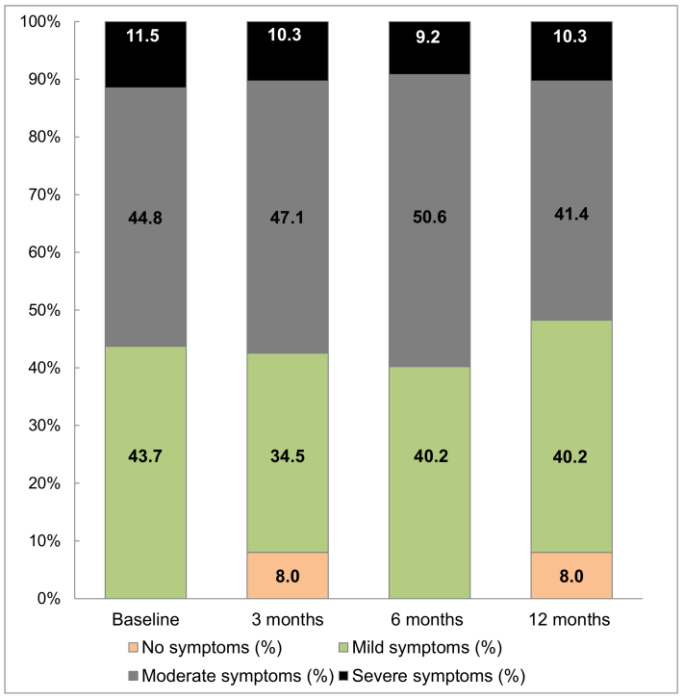


Figure D.4.3 Cluster 3: 3-cluster LCGA anxiety model for complete binary anxiety data (n=56).

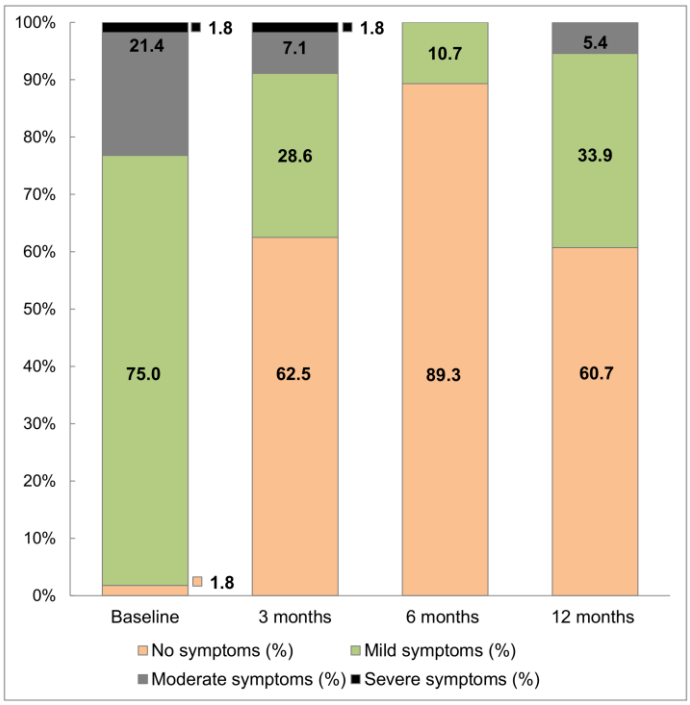
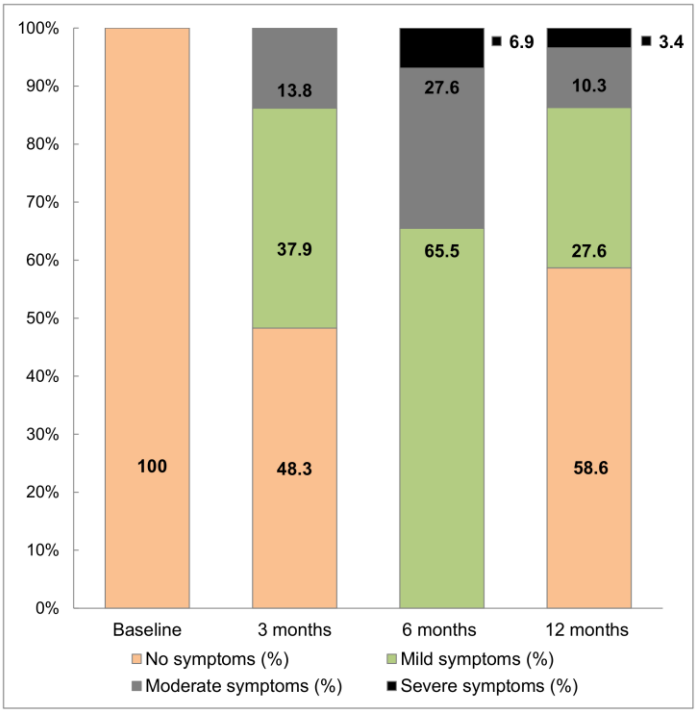


Figure D.4.4 Cluster 4: 3-cluster LCGA anxiety model for complete binary anxiety data (n=29).



D.5 INDIVIDUAL ANXIETY PATTERNS IN 3 CLUSTERS NESTED IN THE 3-CLUSTER LCGA ANXIETY MODEL

Figure D.5.1 Individual HADS-anxiety scores over time for 119 individuals with the *no anxiety symptom trajectory*.

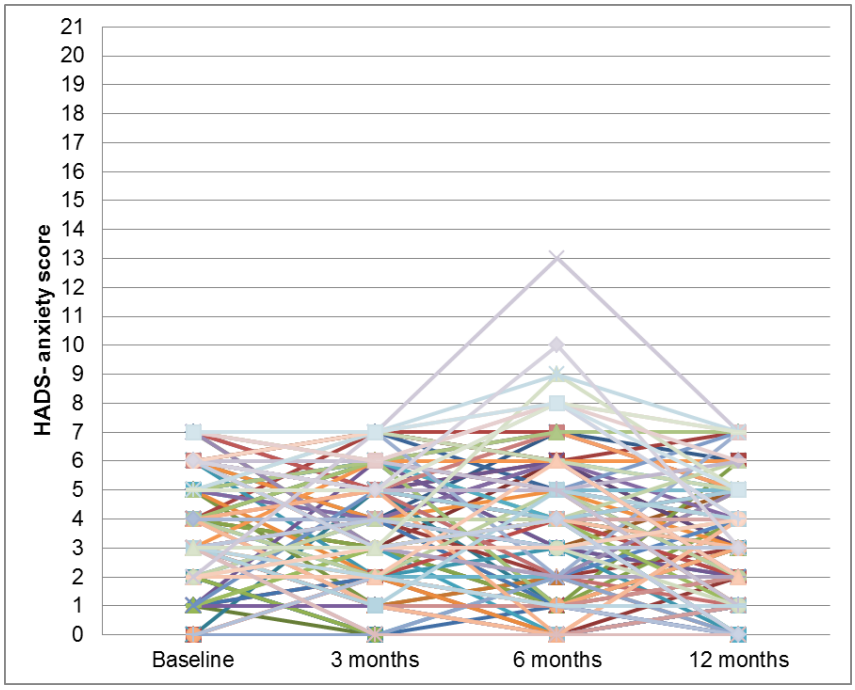


Figure D.5.2 Individual HADS-anxiety scores over time for 86 individuals in the *transient anxiety symptom trajectory*.

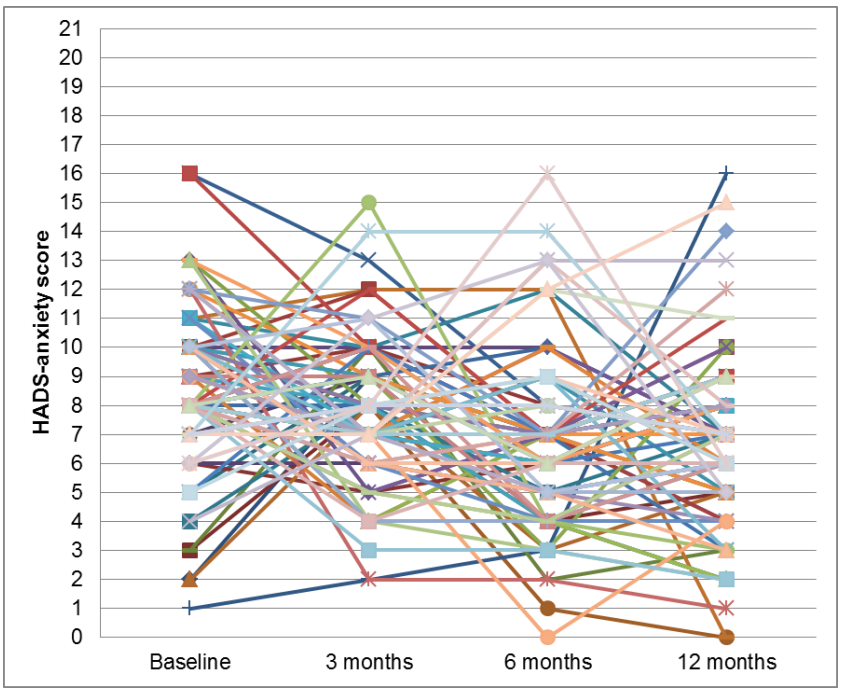
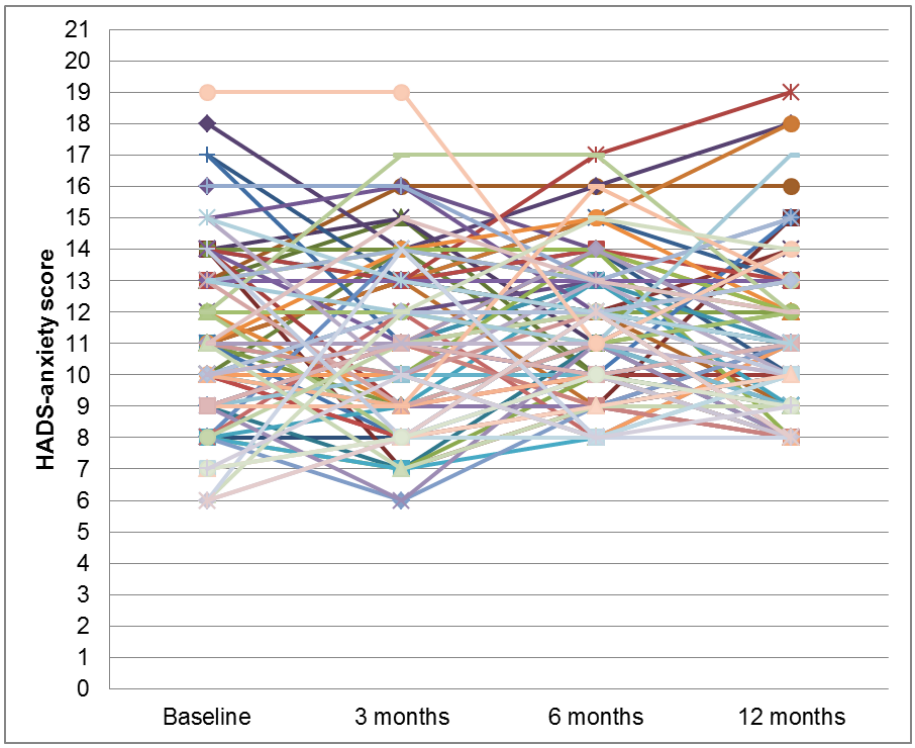


Figure D.5.3 Individual HADS-A scores over time for 88 individuals in the *persistent anxiety symptom trajectory*.



Appendix E: The course of anxiety and depression symptoms in older patients presenting to general practice with musculoskeletal pain

Part 3: Factors associated with the course of anxiety and depression symptoms

This appendix supports selected analyses described in chapter six

E.1 RESULTS OF THE MULTICOLLINEARITY TEST

Table E.1.1 Results of VIF test for multicollinearity.

Variable	VIF	1/VIF
Pain interference with daily activities	5.19	0.19
Pain interference with work	3.83	0.26
Pain interference with social activities	3.70	0.27
Lack of partner	2.83	0.35
Living alone	2.79	0.36
Often using coping self-statements	1.48	0.68
Age 70+	1.37	0.73
Age 60-69	1.32	0.76
Lack of emotional support	1.31	0.76
Often using increased behavioural activities	1.30	0.77
Often ignoring sensations	1.25	0.80
Often catastrophising	1.25	0.80
Pain extent	1.19	0.84
Lack of instrumental support	1.12	0.89
Gender	1.08	0.93
Manual/routine work	1.06	0.94
Mean VIF	2.05	

Note: 1/VIF- tolerance; VIF-variance inflation factor

E.2 SATURATED REGRESSION ANALYSES

Table E.2.1 Saturated logistic regression analysis for the 2-cluster LCGA depression model.

No depression symptoms^ vs. Persistent depression symptoms	Odds Ratio	P	95% CI	
Age 60-69	1.60	0.269	0.70	3.69
Age 70+	3.63	0.002	1.61	8.17
Gender	0.97	0.938	0.50	1.90
Number of pain sites (0-44)	1.07	0.001	1.03	1.12
Interference with daily activities (0-10)	0.81	0.144	0.61	1.07
Interference with social activities (0-10)	1.38	0.006	1.10	1.73
Interference with work (0-10)	1.16	0.183	0.93	1.44
Catastrophising	1.76	0.127	0.85	3.64
Ignoring pain sensations	1.15	0.736	0.51	2.56
Coping self-statements	0.85	0.686	0.39	1.86
Increased behavioural activities	0.56	0.122	0.27	1.17
Lack of partner	1.12	0.851	0.33	3.80
Living alone	1.39	0.627	0.37	5.28
No emotional support	3.89	0.190	0.51	29.73
No instrumental support	2.37	0.161	0.71	7.70
Manual/routine work	0.86	0.673	0.43	1.74

Note: ^- reference group.

Table E.2.2 Saturated multinomial logistic regression analysis for the 3-cluster LCGA anxiety model.

No anxiety symptoms^ vs.	OR	P	95% CI	
Persistent anxiety symptoms				
Age 60-69	0.48	0.054	0.23	1.01
Age 70+	0.70	0.388	0.31	1.57
Gender	1.20	0.578	0.63	2.31
Number of pain sites (0-44)	1.10	0.000	1.05	1.16
Interference with daily activities (0-10)	1.22	0.147	0.93	1.58
Interference with social activities (0-10)	0.92	0.426	0.75	1.13
Interference with work (0-10)	1.09	0.394	0.89	1.34
Catastrophising	3.58	0.002	1.57	8.13
Ignoring pain sensations	1.29	0.530	0.59	2.81
Coping self-statements	0.96	0.926	0.45	2.05
Increased behavioural activities	0.39	0.011	0.19	0.80
Lack of partner	1.75	0.395	0.48	6.37
Living alone	0.68	0.600	0.16	2.84
No emotional support	1.93	0.628	0.14	27.38
No instrumental support	5.74	0.023	1.27	15.99
Manual/routine work	1.33	0.428	0.66	2.66
Transient anxiety symptoms				
Age 60-69	1.26	0.527	0.62	2.57
Age 70+	1.58	0.256	0.72	3.47
Gender	1.22	0.524	0.66	2.28
Number of pain sites (0-44)	1.06	0.019	1.01	1.12
Interference with daily activities (0-10)	1.20	0.154	0.93	1.55
Interference with social activities (0-10)	0.92	0.429	0.76	1.12
Interference with work (0-10)	1.10	0.327	0.91	1.34
Catastrophising	2.77	0.013	1.23	6.20
Ignoring pain sensations	1.22	0.599	0.58	2.57
Coping self-statements	1.38	0.387	0.67	2.82
Increased behavioural activities	0.44	0.020	0.22	0.88
Lack of partner	1.59	0.450	0.48	5.30
Living alone	0.66	0.544	0.17	2.54
No emotional support	1.34	0.838	0.08	21.67
No instrumental support	2.60	0.235	0.54	7.54
Manual/routine work	2.38	0.008	1.26	4.52

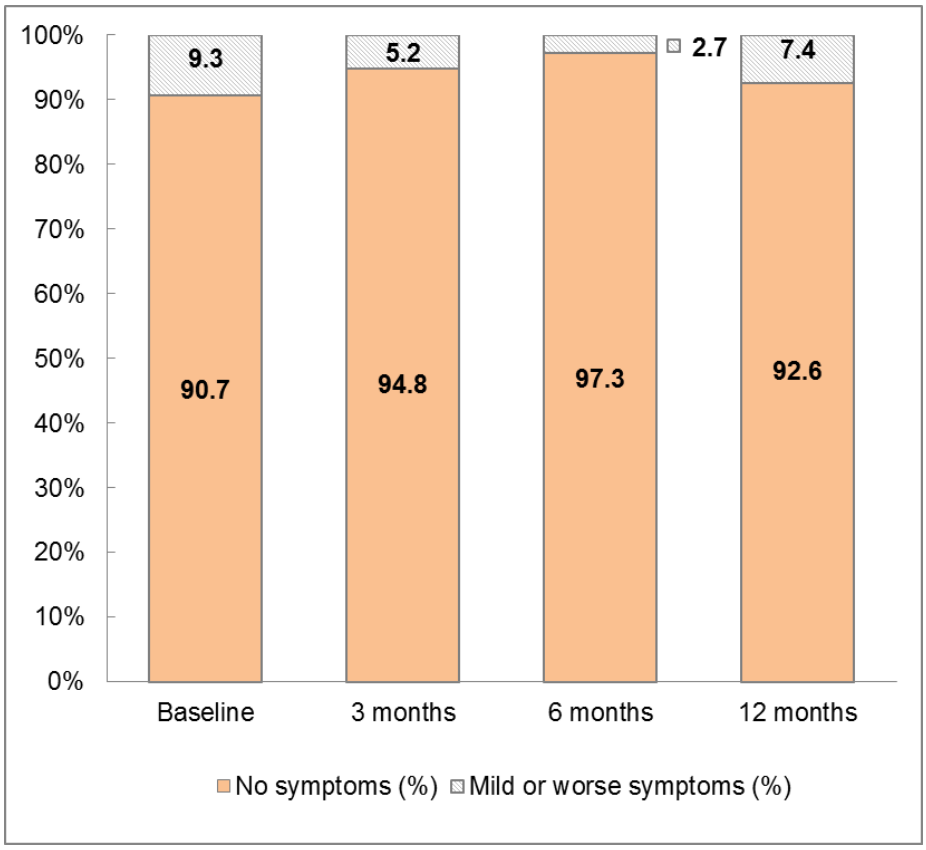
Note: ^- reference group; OR- odds ratio.

E.3 THE 2-CLUSTER LCGA DEPRESSION MODEL BASED ON 368 PARTICIPANTS WITH AT LEAST THREE HADS SCORES AVAILABLE

Table E.3.1 Average assignment probabilities based on maximum posterior probability for 2 clusters LCGA depression model (n=368).

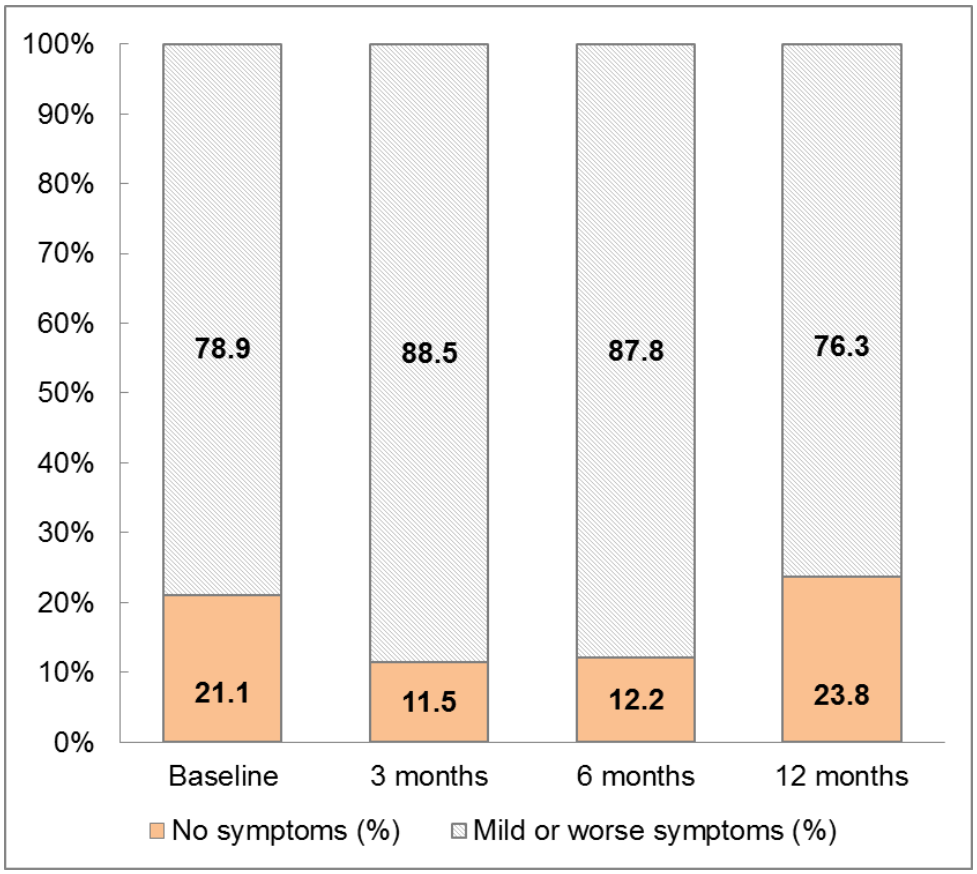
Assigned cluster	n	%	Average posterior probabilities for each cluster	
			1	2
1	272	73.9	**.9877**	.0123
2	96	26.1	.0690	**.9310**

Figure E.3.1 2-cluster LCGA depression model: *no depression symptom trajectory* (n=272).



Note: Frequencies exclude missing data.

Figure E.3.2 2-cluster LCGA depression model: *persistent depression symptom trajectory* (n=96).



Note: Frequencies exclude missing data.

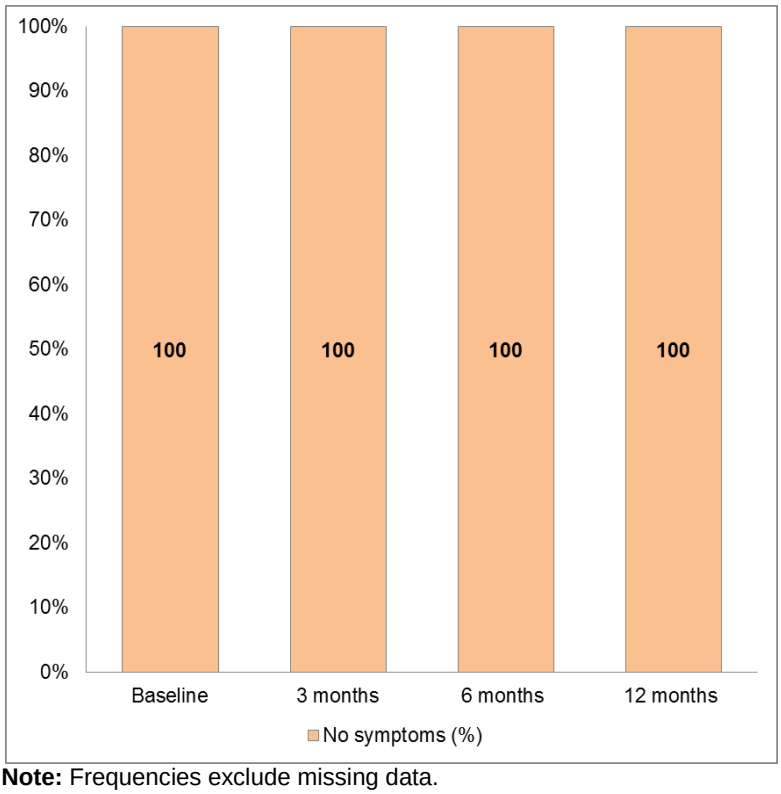
E.4 THE 3-CLUSTER LCGA ANXIETY MODEL BASED ON 368 PARTICIPANTS WITH AT LEAST THREE HADS SCORES AVAILABLE

Table E.4.1 Average assignment probabilities based on maximum posterior probability for 3 clusters LCGA anxiety model (n= 368).

Assigned cluster	n	%	Average posterior probabilities for each cluster		
			1	2	3
1	139	37.8	**.8756**	.0001	.1243
2	123	33.4	.0001	**.8869**	.1130
3	106	28.8	.0853	.0942	**.8205**

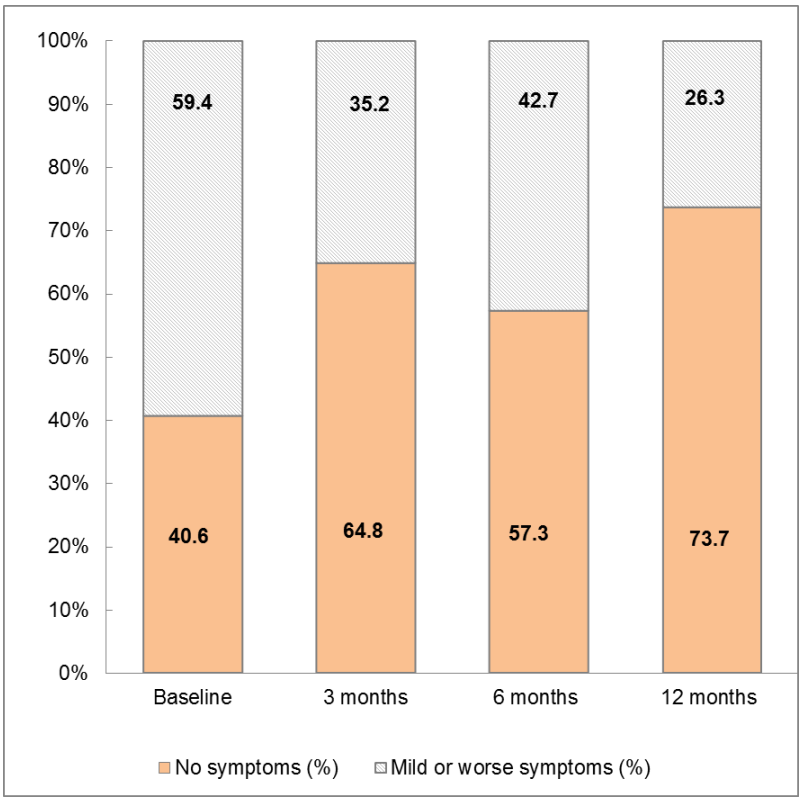
Note: Cluster 1- no anxiety symptoms, Cluster 2- 'transient anxiety symptoms', Cluster 3- 'persistent anxiety symptoms'.

Figure E.4.1 3-cluster LCGA anxiety model: *no anxiety symptom trajectory* (n=139).



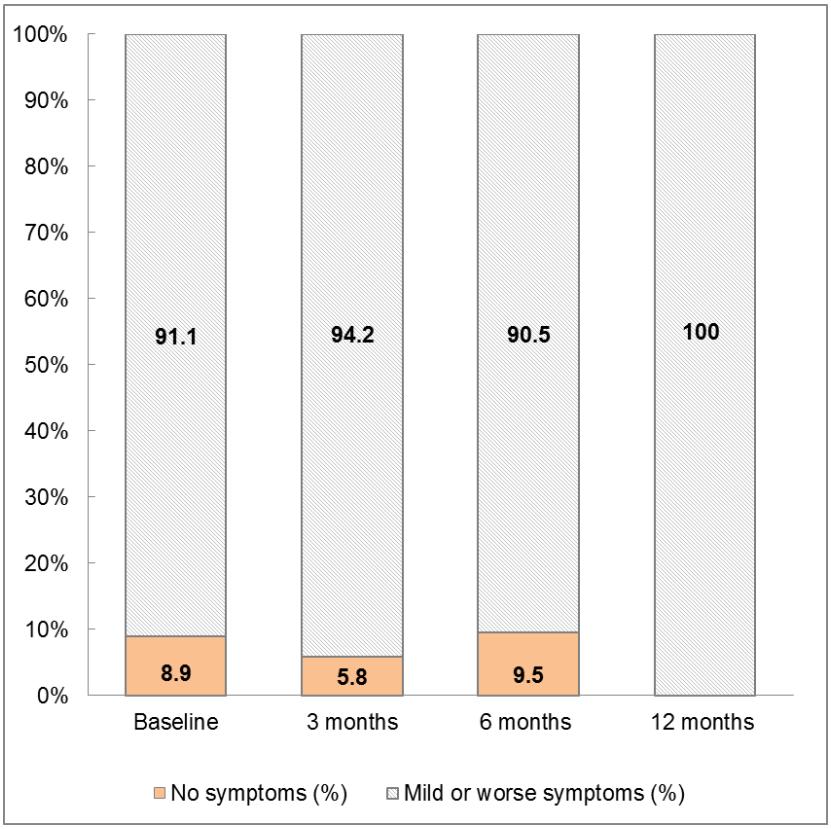
Note: Frequencies exclude missing data.

Figure E.4.2 3-cluster LCGA anxiety model: *transient anxiety symptom trajectory* (n=72).



Note: Frequencies exclude missing data.

Figure E.4.3 3-cluster LCGA anxiety model: *persistent anxiety symptom trajectory* (n=118).



Note: Frequencies exclude missing data.

Appendix F: Documented detection of depression and anxiety in older adults consulting with musculoskeletal pain: analyses of medical record data

This appendix consists of Read terms and codes used in chapter seven

F.1 LIST OF SEARCH TERMS

Table F.1.1 Read term list for depression.

Read code	Read terms
Depression *diagnosis/problem codes*	
1BT..	Depressed mood/ Low mood/Depression
E11z2	Masked depression
Eu32z	[X]Depression NOS
E2B1.	Chronic depression
212S.	Depression resolved
2257.	O/E - depressed
E135.	Agitated depression
1B1U.	Depression symptoms/ Symptoms of depression
E112.	Single major depressive episode/ Major depression Mild major depression/ Moderate major depression
E113.	Recurrent major depressive episode
Eu32.	[X]Depressive episode
Eu341	[X]Depressive personality disorder
E2B..	Depressive disorder NEC
1B17.	C/O - feeling depressed/depressed
E290z	Brief depressive reaction NOS
Eu32z	[X]Depressive disorder NOS
Eu412	[X]Mild anxiety depression
E291.	Prolonged depressive reaction
E130.	Reactive depressive psychosis
Eu33.	[X]Seasonal depressive disorder
Eu33.	[X]Recurrent depressive disorder
E113.	Endogenous depression - recurrent
Eu32z	[X]Depressive episode, unspecified
E112.	Endogenous depression first episode
Eu3y1	[X]Recurrent brief depressive episodes
Eu31.	[X]Bipolar affective disorder
9HA0.	On depression register
9HA1.	Removed from depression register
6891.	Depression screen
388K.	Geriatric depression scale
ZV790	[V]Screening for depression
388P.	HAD scale: depression score
388g.	Beck depression inventory second edition score
388Z.	Depression anxiety stress scales depression score
6896.	Depression screening using questions
388f.	Pt Health Questionnaire score

Table F.1.1 cont. Read term list for depression.

Read code	Read terms
***Interventions*: Antidepressants**	
da...	**OTHER ANTIDEPRESSANT DRUGS:**
da1..	FLUPENTIXOL [ANTIDEPRESSANT]
da2..	TRYPTOPHAN
da3..	FLUVOXAMINE MALEATE
da4..	FLUOXETINE HYDROCHLORIDE
da5..	SERTRALINE HYDROCHLORIDE
da6..	PAROXETINE HYDROCHLORIDE (SEROXAT)
da7..	VENLAFAXINE
da8..	NEFAZODONE
da9..	CITALOPRAM
daA..	REBOXETINE
daB..	MIRTAZAPINE
daC..	ESCITALOPRAM
daD..	AGOMELATINE
d7...	**TRICYCLIC ANTIDEPRESSANTS:**
d71..	AMITRIPTYLINE HYDROCHLORIDE
d72..	[ANTIDEPRESSANT]
d73..	*BUTRIPTYLINE
d74..	CLOMIPRAMINE HYDROCHLORIDE
d75..	DESIPRAMINE HYDROCHLORIDE
d76..	DOSULEPIN HYDROCHLORIDE
d77..	DOXEPIN
d78..	IMIPRAMINE HYDROCHLORIDE [ANTIDEPRESSANT] IPRINDOLE
d79..	LOFEPRAMINE
d7a..	MAPROTILINE HYDROCHLORIDE
d7b..	MIANSERIN HYDROCHLORIDE
d7c..	NORTRIPTYLINE
d7d..	PROTRIPTYLINE HYDROCHLORIDE
d7e..	TRAZODONE HYDROCHLORIDE
d7f..	TRIMIPRAMINE
d7g..	VILOXAZINE HYDROCHLORIDE
d7h..	AMOXAPINE
d8...	**MONOAMINE-OXIDASE INHIBITORS:**
d81..	PHENELZINE
d82..	*IPRONIAZID
d83..	ISOCARBOXAZID
d84..	TRANYLCYPROMINE
d85..	MOCLOBEMIDE
d9...	**COMPOUND ANTIDEPRESSANT DRUGS:**
d911.	*LIMBITROL 5 capsules
d912.	*LIMBITROL 10 capsules
d913.	*MOTIPRESS tablets x28CP
d914.	*MOTIVAL tablets
d915.	*PARSTELIN tablets
d916.	TRIPTAFEN tablets
d917.	TRIPTAFEN-M tablets
d6...	**LITHIUM SALTS:**
d61..	LITHIUM CARBONATE
d62..	LITHIUM CITRATE

Note: *- discontinued

Table F.1.2 Read term list for anxiety.

Read codes	Read terms
Anxiety *diagnosis/problem codes*	
1B13.	Anxious /Anxiety-symptoms/ Anxiousness/ Feeling anxious Level of anxiety/ Anxiety and fear/ Nervously anxious / Acknowledging anxiety/Anxiety about treatment/ Adjustment reaction with anxious mood
E200.	**Anxiety states:**
E2000	Anxiety state unspecified
E2001	Panic disorder
E2002	Generalised anxiety disorder
E2003	Anxiety with depression
E2004	Chronic anxiety
E2005	Recurrent anxiety
E200z	Anxiety state NOS
E202.	**Phobic anxiety:**
E2020	Phobia unspecified
E2021	Agoraphobia with panic attacks
E2022	Agoraphobia without mention of panic attacks
E2023	Social phobia, fear of eating in public
E2024	Social phobia, fear of public speaking
E2025	Social phobia, fear of public washing
E2026	Acrophobia
E2027	Animal phobia
E2028	Claustrophobia
E2029	Fear of crowds
E202A	Fear of flying
E202B	Cancer phobia
E202C	Dental phobia
E202D	Fear of death
E202z	Phobic disorder NOS
Eu515	[X]Dream anxiety disorder
Eu41.	**[X]Other anxiety disorders:**
Eu410	[X]Panic disorder [episodic paroxysmal anxiety]
Eu411	[X]Generalized anxiety disorder
Eu412	[X]Mixed anxiety and depressive disorder
Eu413	[X]Other mixed anxiety disorders
Eu41y	[X]Other specified anxiety disorders
Eu41z	[X]Anxiety disorder, unspecified
Eu40.	**[X]Phobic anxiety disorders**
Eu400	[X]Agoraphobia
Eu401	[X]Social phobias
Eu402	[X]Specific (isolated) phobias
Eu403	[X]Needle phobia
Eu40y	[X]Other phobic anxiety disorders
Eu40z	[X]Phobic anxiety disorder, unspecified
Eu054	[X]Organic anxiety disorder
Eu411	[X]Generalized anxiety disorder
Eu341	[X]Persistent anxiety depression
Eu606	[X]Anxious [avoidant] personality disorder
388b.	Depression anxiety stress scales anxiety score
388N.	HAD scale: anxiety score

Table F.1.2 cont. Read term list for anxiety.

Read code	Read terms
***Interventions*: Anxiolytic medication**	
d2...	**ANXIOLYTICS:**
d21..	DIAZEPAM [ANXIOLYTIC]
d22..	ALPRAZOLAM
d23..	BROMAZEPAM
d24..	CHLORDIAZEPOXIDE
d25..	CHLORMEZANONE
d26..	CLOBAZAM
d27..	CLORAZEPATE DIPOTASSIUM
d28..	HYDROXYZINE HCL [ANXIOLYTIC]
d29..	*KETAZOLAM
d2a..	LORAZEPAM [ANXIOLYTIC]
d2b..	*MEDAZEPAM
d2c..	MEPROBAMATE
d2d..	OXAZEPAM
d2e..	*PRAZEPAM
d2f..	BUSPIRONE HYDROCHLORIDE
d2g..	FLUMAZENIL
gde..	DULOXETINE (SNRIs)
bd1..	PROPRANOLOL HYDROCHLORIDE

Note: *- discontinued

Table F.1.3 Read term list for specialist mental health care (*inteventions*).

Read codes	Read terms
8G94.	Anxiety management training
8BK0.	Depression management programme
8HHq.	Referral for guided self-help for depression
8CAa.	Patient given advice about management of depression
8CQ..	Mental health crisis plan/Agreement of mental health crisis plan/ Completion of mental health crisis plan/Mental health crisis plan discussed with service user/ Mental health care programme approach crisis plan
8Cd..	Further patient given advice
8CX..	Common mental health conditions stepped care model (Care Services Improvement Partnership 2006)
8H75.	Refer to social worker
8H78.	Refer to counsellor
8H7A.	Refer to mental health worker
8H7B.	Refer to community psych. nurse
8H7J.	Refer to occupational therap.
8H7T.	Refer to psychologist
8G...	Psychotherapy/sociotherapy/ Seen by psychiatric nurse/ Community-based occupational therapy service/ Seen by community occupational therapy – service
8HlB.	Urgent referral to psychiatrist
8HHp.	Referral for guided self-help for anxiety
8H49.	Psychiatric referral
8HlK.	Referral for cognitive behavioural therapy
8A2..	Psychiatric monitoring
8H34.	Psychiatric day care
8HK9.	Psychiatric D.V. requested
665..	Psychiatric disorder monitoring
8H23.	Admit psychiatric emergency
8HVO.	Private referral to psychiatrist

F.2 DISTRIBUTION OF THE NUMEBER OF ALL CONSULTATIONS IN 143 PARTICIPANTS ELIGIBLE FOR ANALYSES OF DETECTION

Figure F.2.1 Distribution of the number of consultations in 143 participants eligible for analyses of detection.