RHEUMATOLOGY

Rheumatology 2013;52:2136–2148
doi:10.1093/rheumatology/ket169
Advance Access publication 3 September 2013

Meta-analysis

The prevalence of depression in rheumatoid arthritis: a systematic review and meta-analysis

Faith Matcham[1], Lauren Rayner[1], Sophia Steer[2] and Matthew Hotopf[1]

Abstract

Objective. There is substantial uncertainty regarding the prevalence of depression in RA. We conducted a systematic review aiming to describe the prevalence of depression in RA.

Methods. Web of Science, PsycINFO, CINAHL, Embase, Medline and PubMed were searched for cross-sectional studies reporting a prevalence estimate for depression in adult RA patients. Studies were reviewed in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines and a meta-analysis was performed.

Results. A total of 72 studies, including 13 189 patients, were eligible for inclusion in the review. Forty-three methods of defining depression were reported. Meta-analyses revealed the prevalence of major depressive disorder to be 16.8% (95% CI 10%, 24%). According to the PHQ-9, the prevalence of depression was 38.8% (95% CI 34%, 43%), and prevalence levels according to the HADS with thresholds of 8 and 11 were 34.2% (95% CI 25%, 44%) and 14.8% (95% CI 12%, 18%), respectively. The main influence on depression prevalence was the mean age of the sample.

Conclusion. Depression is highly prevalent in RA and associated with poorer RA outcomes. This suggests that optimal care of RA patients may include the detection and management of depression.

Key words: depression, rheumatoid arthritis, prevalence, meta-analysis, systematic review.

META-ANALYSIS

Introduction

Depression is more common in RA than in the general population [1] and has been associated with increased pain [2], fatigue [3], reduced health-related quality of life [4], increased levels of physical disability [5] and increased health care costs [6]. Depressed RA patients have poorer long-term outcomes, including increased pain [7], more comorbidities [8] and increased mortality levels [9]. Depression may therefore be a useful target for interventions aimed at improving subjective health and quality of life in RA patients. However, prevalence estimates for depression in RA range between 9.5% [10] and 41.5% [11], making it difficult to establish the likely impact of depression in this patient group.

[1]Department of Psychological Medicine, Institute of Psychiatry, King's College London, London, UK and [2]Department of Academic Rheumatology, King's College Hospital, London, UK.

Submitted 28 September 2012; revised version accepted 21 March 2013.

Correspondence to: Faith Matcham, Department of Psychological Medicine, Institute of Psychiatry, King's College London, Weston Education Centre, 10 Cutcombe Road, London SE5 9RJ, UK.
E-mail: faith.matcham@kcl.ac.uk

There are various reasons why this variation in prevalence estimates may exist. First, the term depression is not clear-cut. Making sense of depressive symptoms in the context of chronic physical disease is challenging—it may be difficult to distinguish between patients with a depressive disorder, as opposed to those demonstrating a normal reaction to living with a chronic, debilitating condition. Further, a number of somatic symptoms of depression (e.g. fatigue, poor sleep and loss of appetite) might be expected to occur in RA as part of the disease process. To overcome this, researchers have adapted diagnostic thresholds to define caseness [12] or removed items that may be confounded by RA symptoms, for example, items assessing fatigue or sleep quality [13]. Such variations in definitions of depression may influence prevalence estimates.

Second, there are a multitude of methods available to detect depression. The gold standard method is psychiatric interview and diagnosis according to Diagnostic and Statistical Manual (DSM) [14] or International Classification of Diseases (ICD) [15] criteria. However, such interviews are time consuming and expensive and therefore often not ideal for examining patients in a busy hospital environment [16]. Alternatively, self-report screening

questionnaires, such as the Patient Health Questionnaire (PHQ) [17] and the Hospital Anxiety and Depression Scale (HADS), may be used. These self-report tools are quick and easy to complete, meaning they are often preferred by researchers attempting to collect a large amount of data from a large sample; they are also cheaper to use than diagnostic interviews. Prevalence estimates according to screening tools are often based on predefined thresholds, which may result in overestimations of prevalence, as screening questionnaires tend to prioritize sensitivity over specificity [16].

Study quality may be a further explanation for the variance in prevalence estimates. Small studies lead to variable and imprecise prevalence estimates. Sampling strategies may influence prevalence estimates, with studies using convenience sampling or low participation rates giving unrepresentative samples that may be healthier than the target population [18]. Furthermore, the population studied can impact prevalence estimates; some studies may include patients with specific disease durations, or those using a particular type of medication, which may impact prevalence levels [19, 20].

There has only been one previous systematic review of depression in RA, which examined the strength of the association between depression and RA [21]. As yet no systematic review has provided pooled prevalence estimates of depression in RA. The present study aims to fill this gap. We aimed (i) to present a pooled prevalence level of depression in RA patients; (ii) to provide a summary of the methods used to define depression in RA and (iii) to explore the impact of study characteristics on prevalence estimates.

Materials and methods

Search strategy and selection criteria

The systematic review protocol and data extraction forms were designed in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA; [22]) by F.M. and L.R. F.M. conducted a systematic search of Web of Science, CINAHL, PsycINFO, Medline, Embase and PubMed, from inception to October 2012. Sample search terms can be found in supplementary Appendix S1, available at *Rheumatology* Online.

Inclusion and exclusion criteria

Studies met the following inclusion criteria: (i) Cross-sectional design, baseline cross-sectional data from a longitudinal study or baseline cross-sectional data from a trial, before group allocation. (ii) Reported a prevalence level for depression using diagnostic criteria, a research diagnostic tool or a validated screening tool (Table 1). (iii) Reported prevalence level as the number of participants meeting predefined criteria for depression, or a percentage from which the number of participants meeting criteria for depression could be calculated. (iv) The sample size was $\geqslant 50$.

Studies were excluded if they: (i) used a selective sample (e.g. intervention trials after group allocation); (ii) used a paediatric sample; (iii) retrospectively reviewed medical records to establish depressive symptomatology.

For the meta-analysis, studies using a screening tool without stating the cut-off threshold used to detect depression were excluded. Table 2 provides a full list of the eligible methods of detecting depression, alongside the numbers of articles utilizing each method and the number of participants assessed.

Data extraction and quality assessment

F.M. conducted the primary data extraction. All articles were examined independently by a second reviewer (L.R.). Inter-rater disagreement was minimal, and any disagreements were resolved through discussion and reexamination of the article in consultation with M.H. When multiple publications spanned the years of longitudinal studies, baseline prevalence levels were reported. A 10-point quality assessment tool (supplementary Appendix S2, available at *Rheumatology* Online) was devised to assess sampling method, sample size, participation rate, criteria used to determine depression and the eligibility criteria for participation in the studies. Articles were scored as follows: 0–3 = low quality; 4–6 = low to medium quality; 7–8 = medium to high quality; 9–10 = high quality.

Outcome measures

Outcomes were major depression, minor depression, depressive/mood/affective disorder, dysthymic disorder or adjustment disorder, defined by diagnostic interview or according to a defined threshold on a screening tool.

Statistical analyses

Data were pooled according to diagnosis of depression or screening tool and threshold used to detect caseness. Heterogeneity was found to be moderately high between studies, and therefore random-effects meta-analyses with 95% CIs were conducted with STATA (version 10.0). Heterogeneity was assessed using I^2, with thresholds of $\geqslant 25\%$, $\geqslant 50\%$ and $\geqslant 75\%$ indicating low, moderate and high heterogeneity, respectively [23].

Sensitivity analyses explored whether prevalence estimates were influenced by study design. Planned sensitivity analyses included the following: exclusion of studies with a participation rate $\leqslant 75\%$, or non-reported participation rate; exclusion of studies not stating a sampling strategy, or using a convenience/non-randomized sampling strategy; exclusion of studies that did not state eligibility criteria for inclusion in the study and exclusion of studies using subsets of patients (for example, a female-only sample or patients with limited disease duration). Subgroup analyses were planned by overall study quality, sample size, country of origin and publication year, if there was more than one study in the subgroup. Spearman's correlation analyses with adjusted r^2 assessed the impact of study variables on prevalence estimates. Funnel plots were produced to explore the possibility of publication bias due to preferential publication of small studies reporting high prevalence estimates; Begg-Mazumdar and Egger's tests of publication bias were also performed.

Table 1 Overview of prevalence studies of mood in RA patients

Study ID	Sampling method[a]	Quality[b]	Sample size	Mean age (s.d.), years	Setting[c]	Criteria for detection of depression (threshold)	Women, %	Country	Prevalence, %
Abdel-Nasser 1998	1	8	60	39.7 (10.9)	1	DSM-III-R	80.0	Egypt	23.3
Alishiri 2008	1	5	411	46.8 (12)	1	HADS (9)	87.3	Iran	23.4
Azad 2008	0	0	86	NS	1	HADS (9)	NS	Pakistan	55.8
Barlow 1999	1	3	102	56.3	1	HADS (8/11)	82.4	UK	HADS $\geqslant$ 8:28.4, HADS $\geqslant$ 11:14.7
Bartlett 2003	1	5	77	57.5	5	CESD (9)	80.5	USA	31.2
Chandarana 1987	1	4	86	56.0	1	HADS (9)	74.0	Canada	19.0
Chaney 1996	1	6	58	52.0 (12.5)	1	IDD for DSM-IV	81.0	USA	14.0
Chang 2007	0	2	509	52.0	NS	HADS (8/11/15)	73.0	USA	HADS $\geqslant$ 8:40.7, HADS $\geqslant$ 11:18.5, HADS $\geqslant$ 15:4.5
Chow 2001	0	0	93	49.6 (12.3)	1	HADS (11)	87.0	Malaysia	17.2
Covic 2006	0	0	134	57.9 (12.2)	1	CESD (16)	77.0	Australia	40.0
Covic 2009	0	0	92	56.3 (13.7)	1	HADS (8/11) CESD (16/19) CESD 13 (9/13)	62.0	UK	HADS $\geqslant$ 8:22.6, HADS $\geqslant$ 11:9.7, CESD $\geqslant$ 16:45.3, CESD $\geqslant$ 19:35.9, CESD-13 $\geqslant$ 9:26.6, CESD-13 $\geqslant$ 13:8.1
Cunningham 2003	0	1	141	59.6 (10.3)	NS	CESD (12)	100.0	USA	13.0
Dirik 2010	0	4	117	48.5 (13.2)	4	HADS (8)	84.6	Turkey	55.6
El-Miedany 2002	1	5	80	41.9 (8.4)	1	ICD-10	88.7	Egypt	66.3
Escalante 2000	1	6	236	55.2	1	CESD (16)	62.0	USA	42.0
Fifield 1992	1	4	988	51.0 (10.0)	1	CESD (16)	78.0	USA	32.0
Frank 1988	1	5	137	58.3 (9.6)	5	DIS for DSM-III	24.1	USA	MDD: 17, dysthymia: 40.7
Frank 1991	1	5	74	55.8	1	IDD for DSM-III	NS	USA	DSM-III: 27, DSM-III-R: 16.2
Goodenow 1990	1	6	194	50.7	1	CESD (16)	100.0	USA	22.7
Hagglund 1989	1	6	52	56.5 (11.9)	1	BDI (10/19/30)	61.5	USA	BDI $\geqslant$ 10:35, BDI $\geqslant$ 19:23, BDI $\geqslant$ 30:20
Hanly 2005	1	2	53	52.0	1	HADS (11)	84.9	Canada	0.04
Hewlett 1995	0	0	50	58.0	1	HADS (8/10)	74.0	UK	HADS $\geqslant$ 8:20 HADS $\geqslant$ 10:00
Hewlett 2002	1	5	93	60.0 (10.8)	1	HADS (11)	64.5	UK	20.4
Hider 2009	1	7	159	56.4 (12.2)	1	HADS (8)	72.0	UK	47.5

(continued)

Table 1 Continued

Study ID	Sampling method[a]	Quality[b]	Sample size	Mean age (s.d.), years	Setting[c]	Criteria for detection of depression (threshold)	Women, %	Country	Prevalence, %
Ho 2011	1	6	100	53.7 (13.6)	1	HADS (11)	75	Singapore	15.0
Ichikawa 1995	0	0	92	53.4 (13.3)	1	SRS (40)	82.6	Japan	48.9
Iriarte 2000	1	4	164	52.0 (12.8)	1	SRS (48)	74	Spain	38.0
Isik 2007	1	4	82	52.3 (11.9)	NS	DSM-IV	84.1	Turkey	41.5
Jacobi 2001	0	5	725	59.0 (14.2)	5	CESD (17)	71	The Netherlands	20.3
Karasu 2002	0	0	71	52.8	4	BDI (not stated)	70.4	Turkey	33.8
Karpouzas 2010	1	4	193	NS	NS	PHQ-9 (10)	NS	USA	36.0
Kasle 2008	0	1	148	56.6 (12.3)	1	CESD (27)	77	USA	7.43
Katz 1994	1	6	726	60.4	1	S-GDS (7)	77	USA	14.0
Kobayashi-Gutierrez 2009	1	3	79	NS	1	CESD (16)	NS	Mexico	26.6
Krug 1997	1	3	77	58.2 (11.4)	1	BDI (10)	22.0	USA	35.0
Lindroth 1994	1	6	78	62.0	1	HADS (10)	83.3	Sweden	25.6
Lok 2010	1	9	200	51.4 (10.5)	1	SCID for DSM-IV	79.0	Hong Kong	Major depression: 9.5, depressive disorder: 1.5, dysthymic disorder: 3.5, adjustment disorder and depression: 0.5
MacKinnon 1998	0	4	143	49.6 (11.2)	1	CESD (16)	74.8	Canada	28.7
Margaretten 2011	1	5	466	54.0 (14.0)	1	PHQ-9 (10)	85.0	USA	37.0
Massardo 2001	0	2	75	Median: 53.0	1	CESD (16)	93.3	Chile	47.0
Mella 2010	1	3	62	51.1 (12.8)	1	HADS (7)	83.9	Brazil	53.2
Mikuls 2003	1	5	98	74.6 (3.8)	2	GDS-5 (2)	100.0	USA	24.5
Mo 2010	1	5	97	NS	1	HADS (11)	NS	UK	2.9
Murphy 1988	1	5	80	Median: 62.0	4	PAS for DSM-III	80.0	UK	12.5
Murphy 1999	1	6	62	Median: 59.5	1	HADS (10)	83.9	UK	17.0
Nas 2011	1	5	421	50.1	1	HADS (7)	82.9	Turkey	75.0
Pastor-Oliver 1998	0	2	221	55.4 (12.4)	5	SRS (48)	84.2	Spain	33.5
Penninx 1996	1	6	210	NS	2	CESD (16)	NS	The Netherlands	41.4
Piergiacomi 1989	1	3	50	51.4 (13.5)	1	CESD (19)	74.0	Italy	42.0
Pincus 1996	0	4	163	61.2 (13.7)	1	HADS (8/11)	72.0	UK	HADS $\geqslant$ 8:23 HADS $\geqslant$ 11:15
Pinheiro 2010	0	2	501	51.0	1	HADS (11)	NS	Brazil	20.6
Plach 2003	0	1	156	59.0 (11)	5	CESD (15)	100.0	USA	35.0
Raspe 1987	0	0	75	49.0	NS	BDI-SF (8)	79.0	Germany	22.0
Revenson 1991	0	3	101	51.0	1	CESD (16)	82.0	USA	36.0
Rivero-Carrera 2011	0	0	113	51.0	1	CESD (16)	89.4	Venezuela	29.0

(continued)

Table 1 Continued

Study ID	Sampling method[a]	Quality[b]	Sample size	Mean age (s.d.), years	Setting[c]	Criteria for detection of depression (threshold)	Women, %	Country	Prevalence, %
Scott 2007	0	2	534	NS	1	HADS (11)	NS	UK	18.0
Sharpe 2001	1	6	53	55.1 (14.1)	1	HADS (7)	70.0	Australia	15.0
Sinclair 2010	0	2	125	57.8 (15.4)	3	CESD (23)	73.6	USA	16.0
Smarr 2000	1	5	426	Median: 62.0	1	CESD (10)	57.0	USA	29.8
Spicer 1998	1	4	461	60.8 (13.3)	3	GDS (5/10)	81.0	USA	GDS $\geqslant$ 5:11 GDS $\geqslant$ 10:2
Takeda 2000	0	4	85	56.0 (11.6)	1	SRS (40)	100.0	Japan	56.5
Taylor-Gjevre 2011	1	2	145	54.2 (15.7)	1	CESD (15)	78.0	Canada	37.2
Tomasevic-Todorovic 2011	0	1	60	49.9 (7.6)	1	BDI (16)	88.3	Serbia	63.33
Treharne 2005	1	3	154	56.3 (15.1)	1	HADS (10)	73.0	UK	16.0
Uguz 2009	1	5	83	49.9 (13.1)	1	SCID for DSM-IV	89.2	Turkey	Major depression: 21.8, dysthymia: 13.3
van Hoogmoed 2010	0	4	228	55.9 (10.8)	1	BDI-pc (4)	63.0	The Netherlands	7.0
Wilkins 2000	0	0	96	52.7	1	CESD (16)	87.1	USA	60.0
Worral 2007	1	2	61	Median: 60.0	1	HADS (11)	77.0	UK	11.5
Wright 1996	0	3	141	57.8	1	CESD (16)	45.0	USA	29.8
Wright 1998	0	3	495	60.0	5	CESD (16)	59.6	USA	30.3
Zamani 2010	0	0	81	NS	1	BDI (not stated)	NS	Iran	22.2
Zaphiropoulos 1974	1	2	50	53.7	4	BDI (15)	72.0	UK	46.0

NS: not stated; PAS: Psychiatric Assessment Schedule; SCID: Structured Clinical Interview for DSM; DIS: Diagnostic Interview Schedule. [a]0: convenience/non-randomized, or undefined sampling strategy, 1: consecutive/randomized sampling strategy. [b]Quality rated out of 10: 0–3: low quality; 4–6: medium quality; 7–8: medium-high quality; 9–10: high quality. [c]1: outpatient, 2: database, 3: panel from longitudinal study, 4: inpatient/outpatient, 5: outpatient/community.

TABLE 2 Methods of detecting depression and summary of prevalence and heterogeneity findings

Tool	Definition/threshold	No. of studies	No. of participants	Prevalence, % (95% CI)	Heterogeneity I^2, %
Diagnostic criteria					
DSM	Major depression	4	480	16.8 (10, 24)	73.4
	Dysthymic disorder	3	420	18.7 (−2, 39)	97.2
	Unspecified depression	2	280	6.4 (−4, 17)	88.1
	Depressive disorder	1	200	1.5	—
	Adjustment disorder and depression	1	200	0.5	—
ICD-10	Unspecified depression	1	80	66.3	—
Screening questionnaires					
Beck Depression Inventory (BDI)	10	2	129	34.9 (27, 43)	0.0
	15	1	50	46.0	—
	16	1	60	63.3	—
	19	1	52	23.0	—
	30	1	52	2.0	—
BDI-SF[a]	8	1	75	22.0	—
BDI-pc[b]	4	1	228	7.0	—
Centre for Epidemiological Studies Depression Scale (CESD)	9	1	77	31.2	—
	10	1	426	29.8	—
	12	1	141	13.0	—
	15	2	301	36.2 (31, 42)	0.0
	16	14	3333	36.0 (32, 40)	83.1
	17	1	725	20.3	—
	19	2	142	37.9 (30, 46)	0.0
	23	1	125	16.0	—
	27	1	148	7.4	—
CESD-13[c]	9	1	92	26.6	—
	13	1	92	8.1	—
Geriatric Depression Scale (GDS)	5	1	461	2.0	—
	10	1	461	11.0	—
S-GDS[d]	7	1	726	14.0	—
GDS-5[e]	2	1	98	24.5	—
Hospital Anxiety and Depression Scale (HADS)	7	3	536	48.0 (9, 87)	98.5
	8	7	1193	34.2 (25, 44)	90.9
	9	3	583	32.1 (14, 50)	94.4
	10	4	344	14.9 (4, 26)	90.9
	11	12	2398	14.8 (12, 18)	74.0
	15	1	509	4.5	—
Inventory to Diagnose Depression (IDD)	DSM-III	1	74	27.0	—
	DSM-III-R	1	74	16.2	—
	DSM-IV	1	58	14.0	—
Patient Health Questionnaire (PHQ-9)	10	2	659	38.8 (34, 43)	19.8
Self-Rating Scale (SRS)	40	2	726	52.6 (52, 60)	1.8
	48	2	98	35.3 (31, 40)	0.0

[a]BDI Short Form; [b]BDI for Primary Care; [c]13-item CES-D; [d]Short GDS; [e]5-item GDS.

Results

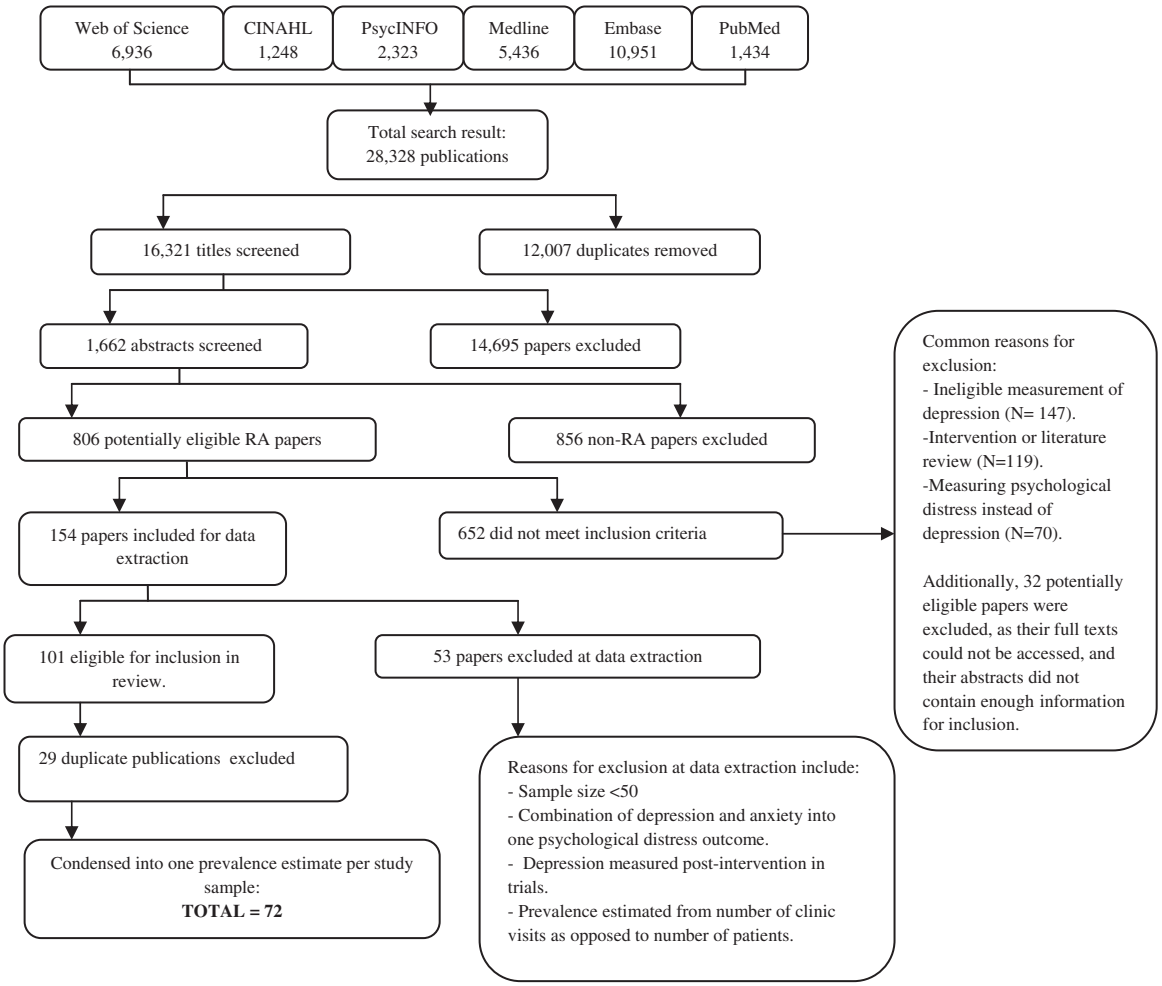
Search results

The search yielded 28 328 relevant articles (Fig. 1). After removal of duplicates, titles and then abstracts were screened for potential eligibility. All non-RA articles were removed, resulting in 806 potentially eligible studies. These were screened according to the inclusion and exclusion criteria for entry into the study, resulting in a total of 101 eligible studies. After taking into account multiple publications from the same sample, 72 articles were included in the review.

Included studies

Table 1 presents the 72 papers included in the review (see supplementary Appendix S3, available at *Rheumatology* Online). Seven studies used diagnostic criteria (DSM or ICD), and the remaining 66 used (one or more) screening

Fig. 1 Search results and study selection.



tools to detect depression (PHQ-9, IDD, HADS, CESD, BDI, SDS or GDS), the most popular being the HADS and the CESD. The studies represented a total of 13 189 patients with RA; the median of mean ages was 53.7 years [interquartile range (IQR) 51.0–56.5], and the median percentage of females represented in the sample was 77.0% (IQR 70.4–82.9%). Sample sizes ranged from 50 to 988 participants (median = 96.0; IQR 75.0–159.0).

Quality assessment

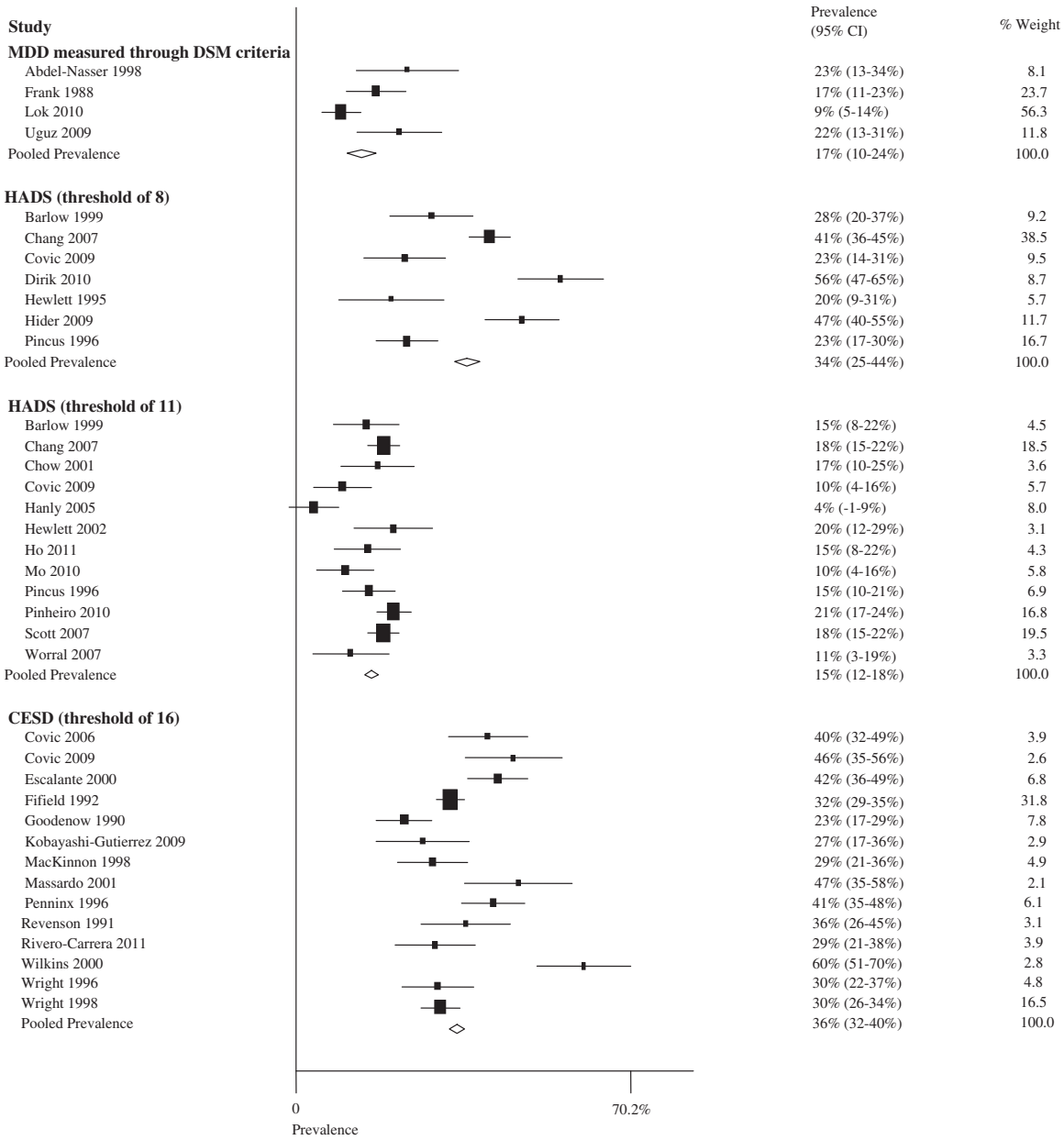
Table 1 presents the quality assessments for the 72 papers, according to the quality assessment tool (supplementary Appendix S2, available at *Rheumatology* Online). The overall quality of the articles was poor with a median quality score of 3/10 (IQR 1–5). Eleven papers (15%) scored 0/10, and 82% of papers scored 5/10 or lower. No papers achieved the maximum score of 10; however, one received 9 out of 10 [10]. Specifically, 16.6% of studies had a sample size larger than 300, only 41.7% stated a participation rate and of these, only 40% had a participation rate $\geqslant$75%. Only 55.6% reported participant eligibility criteria for entry into the study.

Defining depression

Depression was defined in 40 different ways (Table 2). The studies using diagnostic interviews reported three different subtypes of depression: major depressive disorder (MDD), minor depression (MD) and dysthymic disorder (DD), as well as combinations of disorders (depression with adjustment disorders or anxiety) and unspecified depression. Studies using screening questionnaires defined possible or probable caseness using multiple thresholds or detected any depression using one threshold. According to diagnostic criteria, MDD and DD were the most commonly diagnosed depressive subtypes. A full explanation of the differences between depressive diagnoses can be found in supplementary Appendix S4, available at *Rheumatology* Online.

The most commonly used screening questionnaire was the HADS, with 30 studies using this screening tool. However, six different thresholds were presented in the

Fig. 2 Prevalence of MDD in RA.



Pooled prevalence of MDD according to DSM criteria in RA patients by random effects meta-analysis.

articles, with the conventional cut-offs of 8 (probable depression) and 11 (definite depression) being the most commonly used. Twenty-five articles used the CESD; nine different cut-off points were presented, the most commonly used being 16. Eight papers used the BDI, with five different thresholds for depression.

Prevalence of depression

Prevalence of depression alone (excluding combination disorders) ranged between 0.04% and 66.3% in individual studies (Table 1). Table 2 presents the summary of meta-analyses and heterogeneity assessments. Meta-analytical pooled prevalence of MDD (Fig. 2) according to the DSM diagnostic criteria was 16.8% (95% CI 10.0%, 24.0%), with moderate heterogeneity ($I^2 = 73.4\%$). Dysthymic disorder (according to DSM criteria) showed a pooled prevalence level of 18.7% (95% CI −2.0%, 39.0%), with high heterogeneity ($I^2 = 97.2\%$).

Prevalence of depression according to the PHQ-9, with a threshold of 10 indicating moderate-severe depressive symptoms, was 38.8% (95% CI 34.0%, 43.0%), with low heterogeneity ($I^2 = 19.8\%$).

Analyses of screening questionnaires according to the threshold used to detect depression were conducted. As

expected, higher thresholds yielded lower prevalence estimates. For example, the HADS shows an estimated prevalence of 34.2% when used with a threshold of 8, and a prevalence of 14.8% when used with a threshold of 11 (Fig. 2).

Assessment of publication bias (see supplementary Appendix S5, available at *Rheumatology* Online) indicated significant publication bias, according to the Egger's test, in studies reporting MDD according to DSM criteria [Begg-Mazumdar: Kendall's $\tau = 1.36$, $P = 0.17$, Egger: bias = 4.59 (95% CI 1.36%, 7.82%), $P = 0.03$]. There was no significant evidence of publication bias in any other analyses.

Sensitivity and subgroup analyses

Table 3 shows prevalence estimates according to each sensitivity and subgroup analysis, in comparison with the primary analysis. The results of the sensitivity analyses indicated no particular trend or pattern according to the exclusion of studies with only abstracts available, the exclusion of studies with unreported participation rates or participation rates $\leqslant 75\%$, the removal of studies using convenience, non-randomized, or with unreported sampling strategies, or the exclusion of studies using subsets of patients. Exclusion of studies with no reported eligibility criteria tended to increase prevalence estimates, with the exception of the CESD (threshold 16). The subgroup analyses were conducted according to sample size, overall quality and publication year. The subgroup analyses for sample size and overall quality showed no clear patterns. However, more recent publications tended to yield higher prevalence estimates.

Associated study variables

Spearman's correlation analyses with adjusted r^2 were used to assess the associations between linear variables including participation rate, sample size, overall study quality, publication year, proportion of female participants, mean age of participants and mean duration of illness. Table 4 shows the results of these analyses.

A significant relationship was found between mean age and prevalence estimate; lower age was associated with increased depression prevalence ($r = -0.3$, $P = 0.02$). No other study characteristics showed a significant association with prevalence estimate.

Discussion

Depression is highly prevalent in RA patients. Estimates varied according to the way in which depression was measured, but our pooled estimates from the small number of studies using gold standard clinical interviews suggest that major depression is present in 16.8% of RA patients. The larger number of studies using screening tools found significant depressive symptoms present in 38.8% using the PHQ-9 and between 14.8% and 48% using the HADS. These prevalence estimates are considerably higher than those observed in the general population [1] and are similar to, or higher than, those observed in patients with diabetes [24], Parkinson's disease [25] and cancer [26]. Although studies varied widely in terms of quality (and many were of poor quality), our sensitivity analyses indicate that prevalence estimates were reasonably stable. Apart from the measurement tool used to ascertain depression, study quality and study population had little impact on the estimates detected.

The RA patient population represents a largely female, older adult population [27]. It could be suggested that the inflated levels of depression found in this sample represent the increased risk of depression found in females and the elderly [28, 29], regardless of the presence of RA. However, as we found a significant negative association between age and depression prevalence estimate, it is more likely that our findings represent and increased risk of depression in RA patients in comparison with the general population.

A bewildering diversity of assessment measures were used to ascertain depression. This is similar to the situation in other physical diseases [30]. In this review, we did not include many studies that did not use validated measures of depression or questionnaires that assess a broader overlapping concept of psychological distress. Nevertheless we found that many studies used idiosyncratic cut-off scores on screening tools, meaning that the range of estimates for one such measure (the HADS) varied from 14.8% to 48%. Because there have not been validation studies to determine the best cut-point for such screening tools in this population, one clear recommendation is that investigators justify the use of idiosyncratic thresholds, and always report prevalence at conventional cut-points as well, to allow cross-study comparisons.

We used rigorous methods to conduct the review, with a sensitive search, and a reproducible, structured approach to data extraction and synthesis. We took a broadly inclusive approach to inclusion of studies, preferring to include less rigorous studies and explore the impact of study design in sensitivity analyses than to exclude such studies from the outset. It is possible that publication bias affected our results. We explored this using funnel plots and Egger's test where the assumption made was that small studies reporting low prevalence of depression would be less likely to be published than small studies reporting high prevalence. We only found evidence of potential publication bias in the studies that used diagnostic interviews. This is surprising since the efforts taken to conduct such studies are considerable and we would have anticipated these to be least likely to be affected by publication bias.

There are, however, additional important shortcomings in the evidence on prevalence of depression in RA that need to be addressed. The limited number of studies using structured clinical interview and determining depression according to DSM and ICD criteria is a concern. The high rates of depressive symptomatology detected through the screening tools could be due to the overlap between the somatic symptoms of depression and symptoms of RA. Symptoms frequently associated with

Table 3 Impact of study characteristics on prevalence estimates for depression in RA: sensitivity and subgroup analyses

		Depression definition (threshold)							
		Major depression (DSM)	Dysthymic disorder (DSM)	HADS (7)	HADS (8)	HADS (9)	HADS (10)	HADS (11)	CESD (16)
Primary analysis		16.8 (10, 24) I^2 = 73.4% 4 studies 480 patients	18.7 (−2, 39) I^2 = 97.2% 3 studies 420 patients	48.0 (9, 87) I^2 = 98.5% 3 studies 536 patients	34.2 (25, 44) I^2 = 90.9% 7 studies 1193 patients	32.1 (14, 50) I^2 = 94.4% 3 studies 583 patients	14.9 (4, 26) I^2 = 90.9% 4 studies 344 patients	14.8 (12, 18) I^2 = 74.0% 12 studies 2398 patients	36.0 (32, 40) I^2 = 83.1% 14 studies 3333 patients
Sensitivity analyses									
Excluding studies at high risk of bias		—	7.8 (2, 17) I^2 = 83.7% 2 studies 283 patients	—	35.2 (23, 47) I^2 = 83.5% 2 studies, 611 patients	22.4 (18.6, 26.1) I^2 = 2.8% 2 studies 497 patients	—	16.4 (14, 18) I^2 = 16.2% 9 studies 1752 patients	32.9 (30, 38) I^2 = 51.3% 8 studies 2145 patients
Excluding studies with only abstracts available		—	—	65.1 (44, 87) I^2 = 90.8% 2 studies 483 patients	33.0 (21, 45) I^2 = 91.4% 6 studies 684 patients	—	—	13.8 (9, 18) I^2 = 77.3% 7 studies 1137 patients	35.4 (31, 40) I^2 = 83.3% 13 studies 302 patients
Excluding studies with unreported PR or PR <75%		15.5 (2, 29) I^2 = 82.2% 2 studies 260 patients	—	—	41.9 (22, 62) I^2 = 95% 3 studies 440 patients	—	21.6 (14, 29) I^2 = 23.1% 2 studies 140 patients	14.7 (11, 19) I^2 = 24.2% 4 studies 453 patients	37.7 (29, 46) I^2 = 77.6% 3 studies 589 patients
Excluding convenience non-randomized or unreported sampling methods		16.8 (10, 24) I^2 = 73.4% 4 studies 480 patients	18.7 (2, 39) I^2 = 97.2% 3 studies 420 patients	48.0 (9, 87) I^2 = 98.5% 3 studies 536 patients	38.1 (19, 57) I^2 = 90.2% 2 studies 262 patients	—	18.9 (14, 24) I^2 = 25.7% 3 studies 294 patients	12.2 (7, 17) I^2 = 67% 6 studies 506 patients	33.2 (26, 40) I^2 = 85.8% 5 studies 1707 patients
Excluding studies with no reported eligibility criteria for participants		16.8 (10, 24) I^2 = 73.4% 4 studies 480 patients	18.7 (2, 39) I^2 = 97.2% 3 studies 420 patients	48.0 (9, 87) I^2 = 98.5% 3 studies 536 patients	43.8 (29, 59) I^2 = 89.6% 3 studies 379 patients	39.2 (7, 71) I^2 = 96.9% 2 studies 497 patients	21.6 (14, 29) I^2 = 23.1% 2 studies 140 patients	14.9 (10, 20), I^2 = 0% 2 studies 202 patients	28.6 (25, 32) I^2 = 28.5% 6 studies 1153 patients
Excluding studies using subsets of patients[a]		16.8 (10, 24) I^2 = 73.4% 4 studies 480 patients	18.7 (2, 39) I^2 = 97.2% 3 studies 420 patients	—	32.6 (21, 45) I^2 = 92.6% 5 studies 931 patients	—	14.6 (−1, 30) I^2 = 92.2% 3 studies 190 patients	14.7 (12, 18) I^2 = 76.3% 11 studies 2203 patients	37.1 (33, 41) I^2 = 80.7% 13 studies 2932 patients
Subgroup analyses									
Sample size	50–149	19.3 (15, 24) I^2 = 0% 3 studies 280 patients	26.6 (0.3, 53) I^2 = 95.7% 2 studies 220 patients	34.0 (3, 71) I^2 = 95.6% 2 studies 115 patients	30.0 (15, 46) I^2 = 92.6% 4 studies 422 patients	37.1 (0.6, 74) I^2 = 96.7% 2 studies 172 patients	14.6 (−1, 30) I^2 = 92.2% 3 studies 190 patients	12.0 (9, 16) I^2 = 60.3% 8 studies 691 patients	37.8 (31, 45) I^2 = 81.9% 9 studies 974 patients
	150–399	—	—	—	37.0 (24, 50) I^2 = 92.6% 3 studies 832 patients	—	—	—	35 (23, 48) I^2 = 92.2% 3 studies 640 patients
	400+	—	—	—	—	—	—	19.0 (17, 21) I^2 = 0% 3 studies 1544 patients	31.7 (29, 34) I^2 = 0% 2 studies 1483 patients

(continued)

Table 3 Continued

		Depression definition (threshold)							
		Major depression (DSM)	Dysthymic disorder (DSM)	HADS (7)	HADS (8)	HADS (9)	HADS (10)	HADS (11)	CESD (16)
Overall quality	0–3 (low)	–	–	–	28.6 (18, 39) I^2 = 87.3% 4 studies 753 patients	–	8.9 (9, 23) I^2 = 93.7% 2 studies 204 patients	14.6 (12, 19) I^2 = 81.4% 8 studies 1945 patients	37.9 (31, 45) I^2 = 83.2% 8 studies 1326 patients
	4–6 (medium)	18.4 (13, 24) I^2 = 0% 2 studies 220 patients	26.6 (0.3, 53) I^2 = 95.7% 2 studies 220 patients	45.4 (−14, 104) I^2 = 99.2% 2 studies 474 patients	39.3 (8, 71) I^2 = 96.9% 2 studies 280 patients	22.4 (19, 26) I^2 = 2.8% 2 studies 497 patients	21.6 (12, 29) I^2 = 23.1% 2 studies 140 patients	14.7 (11, 19) I^2 = 24.2% 4 studies 446 patients	33.5 (27, 40) I^2 = 85.7% 5 studies 1771 patients
Publication year	1990s	–	–	65.1 (44, 87) I^2 = 90.8% 2 studies 483 patients	24.2 (2, 29) I^2 = 0% 3 studies 315 patients	–	14.6 (−1, 30) I^2 = 92.2% 3 studies 190 patients	15.1 (12, 19) I^2 = 0% 2 studies 265 patients	31.3 (28, 35) I^2 = 69% 7 studies 2272 patients
	2000s	–	–	–	37.3 (25, 49) I^2 = 89.4% 3 studies 761 patients	39.2 (7, 71) I^2 = 96.9% 2 studies 497 patients	–	14.2 (10, 19) I^2 = 81.2% 7 studies 1435 patients	43.5 (35, 52) I^2 = 79.3% 6 studies 712 patients
	post-2010	–	–	–	–	–	–	15.7 (9, 22) I^2 = 77.2% 3 studies 698 patients	–
Country of origin	America	–	–	–	–	–	–	–	35.5 (29, 42) I^2 = 88.7% 7 studies 2251 patients
	UK	–	–	–	28.6 (19, 39) I^2 = 86.4% 5 studies 567 patients	–	11.5 (0.3, 23) I^2 = 90.4% 3 studies 266 patients	14.6 (11, 18) I^2 = 47.9% 7 studies 1142 patients	–
	Asia	–	–	–	–	–	–	16.0 (11, 21) I^2 = 0% 2 studies 193 patients	–

The first line in each set of data is percentage prevalence (95% CI).

[a]Reasons for exclusion: female only sample, limited disease durations examined; only patients using anti-tumour necrosis factor therapy treatment; only one ethnicity represented in sample.

TABLE 4 Spearman's rank correlations between study characteristics and prevalence estimates

Study characteristic	No. of studies	Prevalence estimate	
		ρ	P
Participation rate	30	0.21	0.27
Sample size	72	−0.07	0.53
Overall quality	72	−0.004	0.97
Publication year	72	0.11	0.36
Female, %	64	0.20	0.11
Mean age	60*	−0.30*	0.02*
Mean duration of disease	36	0.02	0.90

*Significant at a $P < 0.05$ level.

depression (such as fatigue and reduced sleep quality) may be experienced by RA patients regardless of whether depressive symptoms are present or not. For example, 7 out of the 21 BDI items assess somatic symptoms, leading to concerns about the validity of this questionnaire in medical patients [31]. Similarly, a modified version of the CESD has been suggested for use with patients with RA, due to the symptom overlap [32]; however, only two articles in the current review used the modified versions available [33, 34].

A further consideration is the representativeness of the sample from which prevalence levels are estimates. Low socio-economic status (SES) patients are often under-represented in research samples [35]. This can be problematic, as low SES is associated with increased susceptibility to depression [36] and RA [37]. Many of the studies included in this review did not measure SES appropriately, with most studies using a single measure of education level or monthly income to indicate SES. This level of heterogeneity makes it difficult to establish the representativeness of the samples included with regard to SES. However, it is possible that a selection bias favouring high SES patients exists and the results of this systematic review may therefore underestimate the prevalence of depression.

The meaning of depression in the context of RA is not straightforward. Emotional responses to a physical illness characterized by pain and debility are understandable, and somatic symptoms of depression (e.g. loss of appetite and poor sleep) might be expected as part of RA. Therefore there is a need to ensure that measures of depression used in clinical practice are validated, both against a recognized criterion (e.g. the 'gold standard' clinical interviews) and also in terms of predictive validity (i.e. to determine the impact of depression on RA outcomes). Psychometric approaches utilizing longitudinal data may further be able to distinguish subtypes of depressive symptoms and thereby distinguish symptoms that are most likely to be core to the depressive syndrome.

Ultimately the key question is whether improved patient outcomes can be attained by recognizing and managing depression more effectively. There is growing evidence that incorporating a system of collaborative and stepped care of depression in patients with physical illness, which might include routine screening for depression with referral for highly structured manualized therapies depending on the outcome of screening, is effective treatment [38]. The high prevalence of depression in RA suggests that this would be a suitable patient group in which to test such strategies.

Rheumatology key messages

- Depression is highly prevalent in RA patient groups.
- Increased depression prevalence in RA is significantly associated with low mean age.

Acknowledgements

We thank Dr William Lee for his statistical advice and Dr Katherine Beck for her assistance with data extraction. F.M., L.R. and M.H. receive salary support from the National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Funding: This work was performed on behalf of the IMPARTS Project: Integrating Mental and Physical Healthcare: Research, Training and Services. IMPARTS is funded by King's Health Partners Academic Health Science Centre, with the overall aim to improve integration of mental and physical healthcare in general hospital settings.

Disclosure statement: The authors have declared no conflicts of interest.

Supplementary data

Supplementary data are available at *Rheumatology* Online.

References

1 Waraich P, Goldner EM, Somers JM *et al*. Prevalence and incidence studies of mood disorders: a systematic review of the literature. Can J Psychiatry 2004;49:124–38.

2 Atal SA, Ceceli E, Okumu M *et al*. The evaluation of pain in patients with rheumatoid arthritis. Pain Pract 2009;9:31.

3 van Hoogmoed D, Fransen J, Bleijenberg G *et al*. Physical and psychosocial correlates of severe fatigue in rheumatoid arthritis. Rheumatology 2010;49:1294–302.

4 Mikuls T, Saag K, Criswell L *et al*. Health related quality of life in women with elderly onset rheumatoid arthritis. J Rheumatol 2003;30:952–7.

5 El-Miedany YM, El Rasheed AH. Is anxiety a more common disorder than depression in rheumatoid arthritis? Joint Bone Spine 2002;69:300–6.

6 Joyce AT, Smith P, Khandker R *et al*. Hidden cost of rheumatoid arthritis (RA): estimating cost of comorbid cardiovascular disease and depression among patients with RA. J Rheumatol 2009;36:743–52.

7 Spicer JG. Health-related quality of life: covariance structural equation modeling in the evaluaton of symptom status and support in persons with rheumatoid arthritis. PhD thesis, University of California, San Francisco 1998:147.

8 Katz PP, Yelin EH. Prevalence and correlates of depressive symptoms among persons with rheumatoid arthritis. J Rheumatol 1993;20:790–6.

9 Ang DC, Choi H, Kroenke K *et al*. Comorbid depression is an independent risk factor for mortality in patients with rheumatoid arthritis. J Rheumatol 2005;32:1013–9.

10 Lok EYC, Mok CC, Cheng CW *et al*. Prevalence and determinants of psychiatric disorders in patients with rheumatoid arthritis. Psychosomatics 2010;51:338–38.e8.

11 Isik A, Koca SS, Ozturk A *et al*. Anxiety and depression in patients with rheumatoid arthritis. Clin Rheumatol 2007; 26:872–8.

12 Geisser ME, Roth RS, Robinson ME. Assessing depression among persons with chronic pain using the Center for Epidemiological Studies Depression Scale and the Beck Depression Inventory: a comparative analysis. Clin J Pain 1997;13:163–70.

13 Covic T, Pallant JF, Tennant A *et al*. Variability in depression prevalence in early rheumatoid arthritis: a comparison of the CES-D and HAD-D Scales. BMC Musculoskelet Disord 2009;10:18.

14 Diagnostic and statistical manual of mental disorders. 4th edn (DSM-IV). Washington, DC: American Psychiatric Association, 1994.

15 ICD-10 classification of mental and behavioural disorders. Diagnostic criteria for research. 10th revision (ICD-10). Geneva: World Health Organisation, 1993.

16 Hotopf M, Chidgey J, Addington-Hall J *et al*. Depression in advanced disease: a systematic review. Part 1. Prevalence and case finding. Palliat Med 2002;16:81–97.

17 Spitzer R, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. JAMA 1999;282:1737–44.

18 Knudsen AK, Hotopf M, Skogen JC *et al*. The health status of nonparticipants in a population-based health study. Am J Epidemiol 2010;172:1306–14.

19 Barlow JH, Cullen LA, Rowe IF. Comparison of knowledge and psychological well-being between patients with a short disease duration (<=1 year) and patients with more established rheumatoid arthritis (>=10 years duration). Patient Educ Couns 1999;38:195–203.

20 Hider SL, Tanveer W, Brownfield A *et al*. Depression in RA patients treated with anti-TNF is common and under-recognized in the rheumatology clinic. Rheumatology 2009;48:1152–4.

21 Dickens C, McGowan L, Clark-Carter D *et al*. Depression in rheumatoid arthritis: a systematic review of the literature with meta-analysis. Psychosom Med 2002;64:52–60.

22 Moher D, Liberati A, Tetzlaff J *et al*. The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009; 62:1006–12.

23 Higgins JPT, Thompson SG, Deeks JJ *et al*. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–60.

24 Barnard KD, Skinner TC, Peveler R. The prevalence of co-morbid depression in adults with Type 1 diabetes: systematic literature review. Diabet Med 2006;23:445–8.

25 Reijnder JSAM, Ehrt U, Weber WEJ *et al*. A systematic review of prevalence studies of depression in Parkinson's disease. Mov Dis 2008;23:183–9.

26 Mitchell AJ, Chan M, Bhatti H *et al*. Prevalence of depression, anxiety, and adjustment disorder in oncological, haematological, and palliative-care settings: a meta-analysis of 94 interview-based studies. Lancet Oncol 2011;12: 160–74.

27 Gabriel SE. The epidemiology of rheumatoid arthritis. Rheum Dis Clin North Am 2001;27:269–81.

28 Piccinelli M, Wilkinson G. Gender differences in depression. Br J Psychiatry 2000;177:486–92.

29 Blazer D, Burchett B, Service C *et al*. The association of age and depression among the elderly: an epidemiologic exploration. J Gerontol 1991;46:M210–5.

30 Wasteson E, Brenne E, Higginson IJ *et al*. Depression assessment and classification in palliative care patients: a systematic literature review. Palliat Med 2009;23: 739–53.

31 Moran PJ, Mohr DC. The validity of the Beck Depression Inventory and Hamilton Rating Scale for Depression items in the assessment of depression among patients with multiple sclerosis. J Behav Med 2005;28:35–41.

32 Martens MP, Parker JC, Smarr KL *et al*. Assessment of depression in rheumatoid arthritis: a modified version of the Center for Epidemiological Studies Depression Scale. Arthritis Rheum 2003;49:549–55.

33 Goodenow C, Reisine ST, Grady KE. Quality of social support and associated social and psychological functioning in women with rheumatoid arthritis. Health Psychol 1990;9:266–84.

34 Wilkins KE. A psychosocial model for the impact of rheumatoid arthritis on well-being. Diss Abstr Int 2000;61: 384.

35 Farmer DF, Jackson SA, Camacho F *et al*. Attitudes of African American and low socioeconomic status white women toward medical research. J Health Care Poor Underserved 2007;18:85–99.

36 Miech RA, Caspi A, Moffitt TE *et al*. Low socioeconomic status and mental disorders: a longitudinal study of selection and causation during young adulthood. Am J Sociol 1999;104:1096–131.

37 Bengtsson C, Nordmark B, Klareskog L *et al*. Socioeconomic status and the risk of developing rheumatoid arthritis: results from the Swedish EIRA study. Ann Rheum Dis 2005;64:1588–94.

38 Katon WJ, von Korff M, Lin EB *et al*. The pathways study: a randomized trial of collaborative care in patients with diabetes and depression. Arch Gen Psychiatry 2004;61: 1042–9.