*Psychological Medicine* (2013), **43**, 1569–1585. © Cambridge University Press 2012
doi:10.1017/S0033291712001717

REVIEW ARTICLE

# Estimating remission from untreated major depression: a systematic review and meta-analysis

**H. A. Whiteford[1,2]*, M. G. Harris[1,2], G. McKeon[1], A. Baxter[1,2], C. Pennell[1], J. J. Barendregt[2] and J. Wang[3]**

[1] *Policy and Evaluation Group, Queensland Centre for Mental Health Research, QLD, Australia*
[2] *School of Population Health, The University of Queensland, QLD, Australia*
[3] *Departments of Psychiatry and Community Health Sciences, University of Calgary, AB, Canada*

**Background.** Few studies have examined spontaneous remission from major depression. This study investigated the proportion of prevalent cases of untreated major depression that will remit without treatment in a year, and whether remission rates vary by disorder severity.

**Method.** Wait-list controlled trials and observational cohort studies published up to 2010 with data describing remission from untreated depression at ⩽2-year follow-up were identified. Remission was defined as rescinded diagnoses or below threshold scores on standardized symptom measures. Nineteen studies were included in a regression model predicting the probability of 12-month remission from untreated depression, using logit transformed remission proportion as the dependent variable. Covariates included age, gender, study type and diagnostic measure.

**Results.** Wait-listed compared to primary-care samples, studies with longer follow-up duration and older adult compared to adult samples were associated with lower probability of remission. Child and adolescent samples were associated with higher probability of remission. Based on adult samples recruited from primary-care settings, the model estimated that 23% of prevalent cases of untreated depression will remit within 3 months, 32% within 6 months and 53% within 12 months.

**Conclusions.** It is undesirable to expect 100% treatment coverage for depression, given many will remit before access to services is feasible. Data were drawn from consenting wait-list and primary-care samples, which potentially over-represented mild-to-moderate cases of depression. Considering reported rates of spontaneous remission, a short untreated period seems defensible for this subpopulation, where judged appropriate by the clinician. Conclusions may not apply to individuals with more severe depression.

*Received 3 May 2012; Revised 27 June 2012; Accepted 5 July 2012; First published online 10 August 2012*

**Key words**: Major depressive disorder, meta-analysis, natural history, spontaneous remission, systematic review.

## Introduction

Depression is a prevalent disorder, and a leading cause of global disease burden (Murray & Lopez, 1997). It causes more disability than any other disorder in high- and middle-income countries (WHO, 2008) and has considerable impact on productivity (Lim *et al*. 2000). Despite the availability of effective pharmacological and psychological interventions, many people remain untreated (Simon *et al*. 2004). It is challenging for policy makers and governments to increase treatment rates cost-effectively. In industrialized nations, 4–7% of the adult population satisfy criteria for major depression in a year (Alonso *et al*. 2004; Kessler *et al*. 2005; Slade *et al*. 2009). But should governments aim for 100% treatment coverage? If not, what is an acceptable population treatment target?

The episodic nature of major depression (Andrews, 2007) necessitates understanding the extent to which recovery will occur without treatment. A meta-analysis of clinical trials by Krøgsboll *et al*. (2009) attributed 35% of improvement in depression severity to spontaneous recovery, and a further 24% to placebo effects. Community-based epidemiological studies suggest that 90–98% of prevalent cases of depression achieve remission within 1 year (McLeod *et al*. 1992; Lewinsohn *et al*. 1994; Kendler *et al*. 1997); however, these estimates include both treated and untreated individuals.

Few studies allow estimation of the proportion of people who will remit from untreated depression (Posternak & Zimmerman, 2000; Posternak *et al*. 2006).

* Address for correspondence: Professor H. A. Whiteford, The University of Queensland, School of Population Health, Queensland Centre for Mental Health Research, Locked Bag 500, Sumner Park BC, QLD 4074, Australia.
(Email: h.whiteford@sph.uq.edu.au)

Rigorous studies of depression using standardized diagnostic and outcome criteria have only been conducted since pharmacological treatments for depression became widespread (Fox, 2002; Posternak *et al*. 2006). Accurate treatment effects could be established by randomized study designs comparing treated and untreated groups (Schorer *et al*. 1968; Ormel *et al*. 1993), but ethical constraints prevent randomizing depressed participants to remain untreated. Placebo-controlled trials are only useful if a 'no-treatment' group allows separation of the disorder's natural history from placebo effects (Musial *et al*. 2007).

Four study types provide a potential basis for inferring spontaneous remission proportions in 'untreated' samples of depressed patients (Posternak *et al*. 2006). Each design has limitations, and because untreated participants are usually not the focus, data reported for these groups are typically less detailed.

The first type comprises longitudinal studies performed before the widespread availability of antidepressants (Hohman, 1938; Schorer *et al*. 1968; Schorer, 1970). In these studies most depressed individuals did not receive treatment; those that did tended to be the most unwell. Inadequate methodology and samples of heterogeneous diagnostic groups limit the validity of these studies (Azorin, 1995). The second type comprises cohort studies of primary-care attenders designed to determine physician detection and treatment rates. Here, it is possible to identify patients who have depression at baseline, and whose disorder goes undetected and/or treatment is not offered during follow-up. These samples may lack generalizability to non-treatment seekers. The third is prospective, observational cohort studies that systematically assess diagnosis and service utilization at multiple time-points. These studies enable untreated samples to be carefully defined and identified. However, treatment is not randomly determined and those who remain untreated are likely to have less severe disorders than treated cases (Coryell *et al*. 1995; Grilo *et al*. 2005). The fourth type comprises randomized controlled trials (RCTs) using a wait-list control group. Here, active treatment is delayed among participants randomly assigned to a wait-list group. Assuming no treatment is sought during the waiting period, this allows the episode to run its natural course. However, intervention outside the study protocol may be uncontrolled.

Posternak & Miller (2001) reviewed psychotherapeutic intervention studies that randomized depressed individuals to wait-list control groups. They found that 19.7% of untreated participants (15/76 from seven studies) remitted over follow-up periods ranging from 4 to 20 weeks. The aim of the current study was to conduct a systematic review and meta-analysis of short-term remission rates for untreated major depression, updating the review by Posternak & Miller (2001) and expanding inclusion criteria to capture further relevant designs. Specifically, we sought to:

(1) predict the probability of remitting from untreated depression over a 12-month period, after adjusting for study-level variables such as follow-up duration; and
(2) examine whether remission rates vary as a function of disorder severity.

## Method

### *Search methodology*

A systematic search conforming with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Moher *et al*. 2009) was conducted. Medline, PsycINFO and EMBASE were searched using OVID from their inception years to September 2010. The following search string was used: depress* AND (longitudinal OR wait* OR prospective OR follow* OR naturalistic OR cohort OR observational) AND (untreat* OR remiss* OR unrecogni*). Searches were restricted to studies of humans classified as journal articles, clinical trials, meta-analyses or reviews.

Reference lists of retrieved articles were searched to identify additional studies. Four authors were contacted to clarify study results. Three authors with known access to datasets with the potential to derive remission estimates were contacted. They were also asked to suggest relevant studies; one provided previously unpublished data.

Initially, titles and abstracts were assessed for relevance. Full-text versions of potentially eligible papers were retrieved. Primary articles from potentially relevant secondary sources such as meta-analyses or reviews were obtained.

### *Inclusion criteria*

Eligible studies fulfilled four criteria:

(1) *Diagnosis of major depression*. Participants experiencing a major depressive episode at study entry were identified by: (i) a structured clinical interview mapping to a major classification system for mental disorders [e.g. DSM, ICD, RDC] or (ii) scores exceeding thresholds on standardized symptom severity instruments such as the Beck Depression Inventory (BDI; Beck & Steer, 1988) or the Hamilton Depression Rating Scale (HAMD; Frank *et al*. 1991), congruent with a DSM or ICD

https://doi.org/10.1017/S0033291712001717 Published online by Cambridge University Press

diagnosis of major depression. Where diagnostic breakdown was reported, ⩾75% of the sample were classified as having major depression.

(2) *Study design*. Study design potentially allowed for the investigation of an untreated sample (Posternak *et al*. 2006); that, is a study reported on a wait-list group or an observational cohort design.

(3) *Definition of remission*. The study applied clearly defined remission criteria based on either a rescinded diagnosis or a below-threshold score on a standardized symptom severity measure (usually a BDI score <10 or a HAMD score ⩽7).

(4) *Remission data*. The study quantified the number or percentage of 'untreated' participants remitted at one or more follow-up time-points up to and including 2 years after study entry. As the focus was on 1-year remission rates, it was determined, *a priori*, that only remission data from follow-up periods up to 2 years following study entry would be included.

### *Exclusions*

Studies targeting samples with other mood disorder diagnoses (e.g. dysthymia, intermittent depression or minor depression) were out of the scope of this study. Studies defining remission exclusively through improved depression scale scores were excluded. Although it is generally accepted that a 50% improvement in symptom severity scores indicates clinical response, significant residual symptoms may persist, making this an inadequate remission definition (Mendlewicz, 2009). Studies using a 'treatment-as-usual' wait-list condition referring participants back to their physician were excluded. These designs were considered to actively encourage treatment, and it could not be ascertained how many participants remained untreated. Samples where more than one-third of untreated participants were known to have received interventions for depression during follow-up were excluded. Studies investigating specific groups (such as refugee populations) that were not considered representative of the general population were also excluded.

### *Data extraction*

Data were extracted independently by two authors (G.M. and C.P.). Recorded fields included: study identifiers (authors, publication year, country); study descriptors (design, sample/setting, diagnostic inclusion criteria, remission criteria, follow-up duration); untreated sample characteristics (study entry treatment restrictions, definition, treatments reported during follow-up); sample demographics (gender and age distribution); and outcomes (number and/or proportion of total sample remitted, number and/or proportion of followed-up sample remitted). Most studies assessed remission status at or within a short period (2 weeks) prior to follow-up. Two studies (Wang, 2004; Posternak *et al*. 2006) assessed the occurrence of remission over longer periods. For consistency, the follow-up time-points for these studies were taken as the mid-point between the last follow-up period and that of the reported remission estimate. Some studies used multiple acceptable remission definitions; all were documented. Extracted data were cross-checked by two authors (M.H. and H.W.). Discrepancies were resolved through discussion and consensus.

### *Assessment of methodological parameters*

Methodological parameters of included studies (sampling method, diagnostic assessment, diagnostic heterogeneity, 'untreated' sample definitions, remission parameters, and follow-up rates) were evaluated using a template derived from existing indices (McGrath *et al*. 2004; Calabria *et al*. 2010) (see online supplementary material).

### *Statistical analysis*

The probability of remission from untreated depression over a 12-month period was estimated using a regression model with logit transformed remission proportion as the dependent variable. For each study, the data closest to 52 weeks were identified, and the sample size, number of cases remitted and follow-up time in weeks were extracted and tabulated with other study characteristics. Where studies provided remission data based on multiple definitions, a rescinded diagnosis was preferred, followed by clinician-rated symptom scales (e.g. the HAMD) and then participant-rated symptom scales (e.g. the BDI). Because one-third of studies determined 'untreated' status at follow-up, rather than at baseline, we used the follow-up data from each study.

As proportions generally follow a binomial distribution, the proportion of remitted cases was logit transformed and weighted by sample size. The follow-up period for each study was centered at 52 weeks. This provides an intercept value that can be exponentiated to give the probability of remission at 12 months. To identify sources of study variability, we explored the effects of participant and study characteristics within the model. These were: age, explored as both a continuous (median age of sample) and a categorical variable (adult, child/adolescent, older

https://doi.org/10.1017/S0033291712001717 Published online by Cambridge University Press

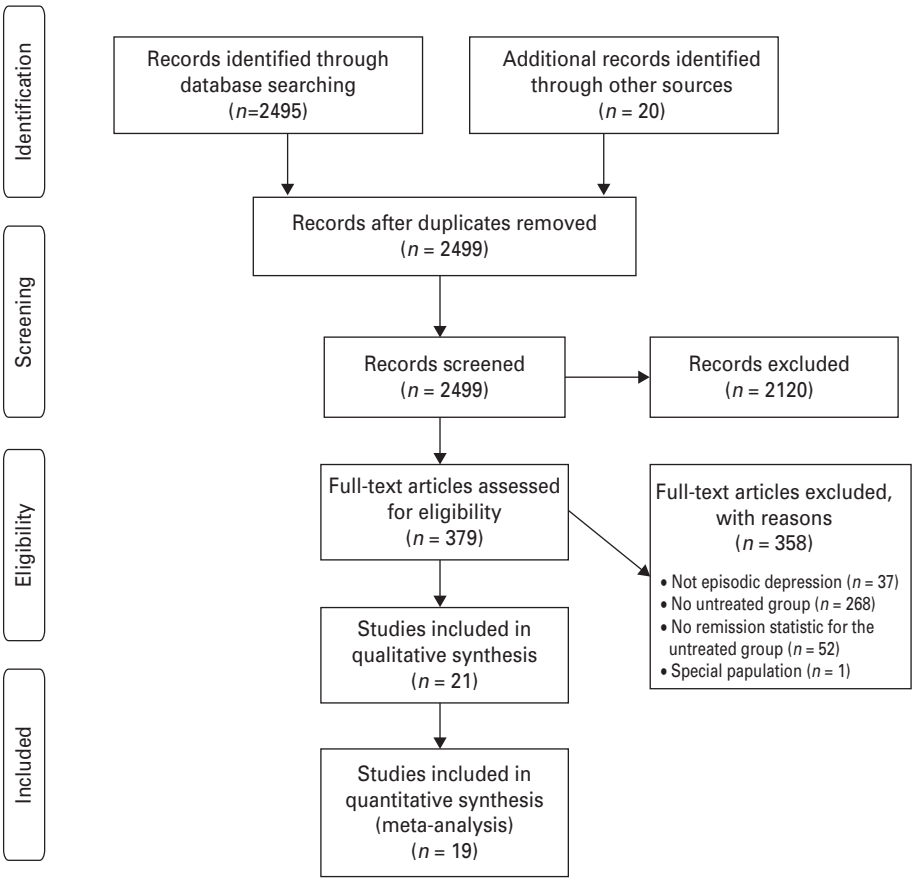


**Fig. 1.** PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart.

adult); gender (percentage female); study type (wait-listed, community-based); and type of diagnostic measure used at follow-up (structured clinical interview *versus* symptom scale). Few studies reported remission according to symptom severity, therefore these data were examined descriptively.

Analyses were conducted in Stata 11.0 (Stata Corporation, USA) using the 'blogit' command, which produces maximum-likelihood logit estimates for grouped data. Robust standard errors were calculated to accommodate possible overdispersion in study estimates. The maximum-likelihood approach is highly sensitive to outliers so we calculated deviance residuals to identify influential studies. The relative impact of outliers on the model was explored using Cook's Distance, using the leave-one-out approach. Influential studies were reviewed to identify factors that might explain the variability in estimates and justify exclusion.

The most informative regression model was identified through backward stepwise regression. The probability of remitting within the follow-up period was estimated using the 'margins' post-estimation command. The average weighted survival for each study sample was calculated using the 'predict' command (Buis, 2007).

## Results

### *Included studies*

Twenty-one studies met inclusion criteria (Fig. 1). The majority of participants were female (mean = 73%). Ages ranged between 13 and approximately 80 years (mean = 34 years). Included studies covered 32 years of research. As the duration of follow-up increased, the number of studies reporting remission data decreased. Key features of the included studies are summarized in the following sections.

### *Wait-list control groups*

Thirteen samples were wait-list groups; see Table 1 for characteristics and Table 2 for remission parameters. Sample sizes were modest (mean = 25 participants, range 8–52). Most (93%) reported remission proportions for follow-up periods of $\leqslant$6 months. Half were conducted in general adult samples, two in older adults ($\geqslant$55 years) and five in children/adolescents.

https://doi.org/10.1017/S0033291712001717 Published online by Cambridge University Press

**Table 1.** *Summary of studies reporting short-term remission from untreated depression*

| | Study information | | | | Untreated sample | | | Demographics | |
|---|---|---|---|---|---|---|---|---|---|
| Study population | Author, year (country) | Study design | Setting/sample selection | Diagnostic inclusion criteria | Sample untreated at baseline? | Definition | Treated elsewhere? | % | Mean age (range) |
| **Wait-list control studies** | | | | | | | | | |
| General adults | Allen *et al.* 1998 (USA) | RCT comparing acupuncture targeting depression *v.* non-specific acupuncture *v.* WLCs | Not described/ newspaper advertisements | SCID (DSM-IV) | Individuals engaged in any current treatment for depression were excluded prior to study entry | Individuals randomized to a wait-list condition did not receive the active intervention during follow-up | N.S. | 100 | −(18–45)[a] |
| | Allen *et al.* 2006 (USA) | RCT comparing acupuncture targeting depression *v.* non-specific acupuncture *v.* WLCs | Not described/ newspaper advertisements | $HAMD_{17} \geqslant 14$, SCID (DSM-IV) | Individuals engaged in any current relevant treatment for depression were excluded prior to study entry | Individuals randomized to a wait-list condition did not receive the active intervention during follow-up | N.S. | 71 | 42 (18–65)[b] |
| | Pace & Dixon, 1993 (USA) | RCT comparing cognitive therapy *v.* no treatment control group | Single university campus/ participation for course credit | BDI ⩾10 ⩽29 | Individuals currently receiving psychological or psychiatric treatment for depression were excluded prior to study entry | Individuals randomized to a no treatment control condition did not receive the active intervention during follow-up | N.S. | 77 | 22 (−) |
| | Selmi *et al.* 1990 (USA) | RCT comparing computerized CBT *v.* standard CBT *v.* WLCs | Not described/ newspaper advertisements | SCL-90-R ⩾65th percentile, BDI ⩾16, SADS (RDC) | Not specified | Individuals randomized to a wait-list condition did not receive either active intervention during follow-up | N.S. | 67 | 31 (−) |
| | Shaw, 1977 (Canada) | RCT comparing cognitive therapy *v.* behavior therapy *v.* non-directive therapy *v.* WLCs | Single university campus/health service referrals | Multiple assessments: 'Clinical interview', BDI ⩾18, $HAMD_{17} \geqslant 20$ | Not specified | Individuals randomized to a wait-list condition did not receive either active intervention during follow-up | N.S. | 75 | 20 (18–25) |
| | Wierzbicki & Bartlett, 1987 (USA) | RCT comparing individual cognitive therapy *v.* group cognitive therapy *v.* WLCs | Not described/local advertisements | 'Structured interview' (DSM-III) and BDI ⩾8 ⩽35 | Individuals currently or previously engaged in psychotherapy or pharmacotherapy for depression were excluded prior to study entry | Individuals randomized to a wait-list condition did not receive either active intervention during follow-up | N.S. | – | − (−) |

**Table 1** (*cont.*)

| | Study information | | | | Untreated sample | | | Demographics | |
|---|---|---|---|---|---|---|---|---|---|
| Study population | Author, year (country) | Study design | Setting/sample selection | Diagnostic inclusion criteria | Sample untreated at baseline? | Definition | Treated elsewhere? | % | Mean age (range) |
| Older adults | Arean *et al.* 1993 (USA) | RCT comparing problem-solving therapy *v.* reminiscence therapy *v.* WLCs | Not described/local advertisements | BDI ⩾20, $GDS_{30}$ ⩾10, $HAMD_{17}$ ⩾18, SADS (RDC) | Individuals engaged in current psychological or pharmacological treatment for depression were excluded prior to study entry | Individuals randomized to a wait-list condition did not receive either active intervention during follow-up | N.S. | 75 | 66 (⩾55 ⩽80)[b] |
| | Strachowski *et al.* 2008 (USA) | RCT comparing CBT *v.* WLCs | Single university clinic/local advertisements | BDI >10, DISH (DSM-IV) | Not specified | Individuals randomized to a wait-list condition did not receive the active intervention during follow-up | N.S. | 72 | 62 (⩾55)[b] |
| Children and adolescents | Clarke *et al.* 1999 (USA) | RCT comparing group CBT *v.* group CBT with parent group *v.* WLCs | Two sites/local advertisements, health service referrals | K-SADS-E (DSM-III-R) | Individuals currently engaged in treatment for depression were excluded prior to study entry | Individuals randomized to a wait-list condition did not receive either active intervention during follow-up | Participants were excluded for obtaining non-study depression treatments. Proportion excluded from the untreated sample not specified | 71[a] | 16[a] (14–18)[b] |
| | Diamond *et al.* 2002 (USA) | RCT comparing attachment-based family therapy *v.* WLCs | Not described/school and parent referrals | Two BDIs ⩾16 1 week apart, K-SADS-P (DSM-III-R) | Individuals engaged in psychotherapy or pharmacological treatment for depression were excluded prior to study entry | Individuals randomized to a wait-list condition did not receive the active intervention during follow-up | N.S. | 78[a] | 15[a] (13–17)[a] |
| | Kowalenko *et al.* 2005 (Australia) | RCT comparing CBT *v.* WLCs | Eleven Sydney high schools/ consenting students | CDI >18 | Not specified | Individuals randomized to a wait-list condition did not receive the active intervention during follow-up | N.S. | 100 | 15 (13–16)[b] |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Reynolds & Coats, 1986 (USA) | RCT comparing CBT *v.* relaxation training *v.* WLCs | Single high school/ consenting students | BDI ⩾ 12, BID ⩾ 20, RADS ⩾ 72 (DSM-III) | Individuals currently engaged in pharmacotherapy or other treatments for depression were excluded prior to study entry | Individuals randomized to a wait-list condition did not receive either active intervention during follow-up | N.S. | 63[a] | 16[a] (−) |
| | Weisz *et al*. 1997 (USA) | RCT comparing CBT *v.* no treatment control group | Three elementary schools/consenting students | CDI ⩾ 11, CDRS-R ⩾ 34 | Not specified | Individuals randomized to a no treatment condition did not receive the active intervention during follow-up | N.S. | 47 | 10 (−) |
| **Primary-care physician detection studies** | | | | | | | | | |
| General adults | Goldberg *et al*. 1998 (15 sites, 14 countries) | Prospective cohort study of PC attenders (WHO PPGHC study) | Fifteen PC practices/ consecutive patients | GHQ-12 variable threshold, CIDI-PC (ICD-10, DSM-IV) | At baseline unrecognized patients had not been detected and treated pharmacologically by their PC physician | Subsample not detected and treated pharmacologically by PC physician at baseline | N.S. | 75 | 38 (>17)[b] |
| | Rost *et al*. 1998 (USA) | Prospective cohort study of PC attenders | Arkansas PC practices/ random telephone screening of Arkansas households | Burnam Screener ⩾ 0.06, DIS and ⩾ 5 depressive symptoms in the past 2 weeks | Not specified | Subsample not diagnosed or treated pharmacologically by PC physician during follow-up. Subsample was not referred to specialty mental health care services and did not seek medical help for depression during follow-up | N.A. | 73 | 45 (⩾18)[b] |
| | Schulberg *et al*. 1987 (USA) | Prospective cohort study of PC attenders | Three Pittsburgh PC practices/consenting patients who had received no care in the past 6 months | DIS (DSM-III) | Patients were new to the facility, or received no care in the previous 6 months | Subsample not diagnosed by PC physician during follow-up | Three unrecognized remitted patients were treated at a psychiatric facility, one received an antidepressant and three received ancillary medications during follow-up | 76[a] | 33[a] (−) |
| | Simon *et al*. 1999 (15 sites, 14 countries) | Prospective cohort study of PC attenders (WHO PPGHC study) | Fifteen PC practices/ consecutive patients | GHQ-12 variable threshold, CIDI-PC (ICD-10) | At baseline unrecognized patients had not been detected and diagnosed by their PC physician | Subsample not recognized and diagnosed by PC physician at baseline | N.S. | 73 | 39 (<65)[b] |

**Table 1** (*cont.*)

| Study information | | | | | Untreated sample | | | Demographics | |
|---|---|---|---|---|---|---|---|---|---|
| Study population | Author, year (country) | Study design | Setting/sample selection | Diagnostic inclusion criteria | Sample untreated at baseline? | Definition | Treated elsewhere? | % | Mean age (range) |
| Older adults | Licht-Strunk *et al*. 2009 (The Netherlands) | Prospective cohort study of PC attenders | Fourteen PC practices/ consecutive patients | $GDS_{15}$ >5, PRIME-MD ⩾5 depressed mood and anhedonia (DSM-IV) | At baseline undetected patients were not taking antidepressants, nor had they been referred to a mental health care professional | Subsample not detected or treated by PC physician (pharmacologically or mental health care referral) during follow-up | N.A. | 59 | 64 (>55)[b] |
| **Observational cohort studies in treatment settings** | | | | | | | | | |
| General adults | Posternak *et al*. 2006 (USA) | Prospective cohort study | Five academic medical centers/consenting depressed patients | SADS (RDC) | No somatic treatment had been provided for at least the first 4 weeks of the episode | Subsample did not receive somatic therapy for the entire duration of a depressive episode | N.A. | 66 | 34 (17–74) |
| | Schulberg *et al*. 1997 (USA) | Prospective component of an RCT of treatments for depression | Four Pittsburgh PC practices/presenting patients not receiving treatment for depression | Psychiatrist's assessment, DIS (DSM-III-R) | Individuals engaged in current treatment for depression were excluded prior to study entry | Subsample did not receive depression- specific treatment or referral from PC physician during follow-up | N.A. | 87[a] | 39[a] (18–65)[b] |
| **Community-based epidemiological cohort studies** | | | | | | | | | |
| General adults | Wang, 2004 (Canada) | Longitudinal component of the Canadian National Population Health Survey | Canadian health services/multiple-stage, stratified random sampling procedures | CIDI-SFMD (DSM-IV) | Untreated cases at baseline had not taken antidepressants in the past month, nor sought help for emotional or mental health problems from health professionals in the past 12 months | Subsample had not taken antidepressants in the past month, nor sought help from health professionals for emotional or mental health problems in the past 12 months | N.A. | 66 | 34 (>12)[b] |

BDI, Beck Depression Inventory; BID, Bellevue Index of Depression; CBT, cognitive behavioral therapy; CDI, Children's Depression Inventory; CDRS-R, Children's Depression Rating Scale (revised); CIDI-PC, Composite International Diagnostic Interview – Primary-care version; CIDI-SFMD, Composite International Diagnostic Interview – Short Form for Major Depression; DIS, Diagnostic Interview Schedule; DISH, Depression Interview and Structured Hamilton; F, female; $GDS_{15}$, 15-item Geriatric Depression Scale; $GDS_{30}$, 30-item Geriatric Depression Scale; GHQ-12, 12-item General Health Questionnaire; $HAMD_{17}$, 17-item Hamilton Rating Scale for Depression; K-SADS-E, Schedule for Affective Disorders and Schizophrenia for school-age children – epidemiologic version; K-SADS-P, Schedule for Affective Disorders and Schizophrenia for school-age children – present and lifetime version; N.A., not applicable; N.S., not stated; PC, primary care; PPGHC, psychological problems in general health care; PRIME-MD, Primary-care Evaluation of Mental Disorders; RADS, Reynolds Adolescent Depression Scale; RCT, randomized controlled trial; RDC, Research Diagnostic Criteria; SADS, Schedule for Affective Disorders and Schizophrenia; SCL-90-R, 90-item Hopkins Symptom Checklist (Revised); WHO, World Health Organization; WLCs, wait-list controls; –, not available.

[a] Value not available for untreated subsample, value from total sample provided.

[b] Age range (years) refers to eligibility criteria rather than untreated sample descriptive statistics.

**Table 2.** *Summary of remission parameters*

| Study population | Author, year (country) | Remission definition | Follow-up (weeks) | Proportion remitted | | | |
|---|---|---|---|---|---|---|---|
| | | | | Total sample | | Followed up | |
| | | | | % | *n* | % | *n* |
| **Wait-list control studies** | | | | | | | |
| General adults | Allen *et al*. 1998 (USA) | Absence of core symptoms of depression (depressed mood and anhedonia) (SCID, DSM-IV)[a] | 8 | 18 | 2/11 | 20 | 2/10[b] |
| | Allen *et al*. 2006 (USA) | $HAMD_{17}$ <7 and >50% $HAMD_{17}$ reduction since intake[a] | 8 | 8 | 4/52 | 9 | 4/44 |
| | Pace & Dixon, 1993 (USA) | BDI <10[a] | 5.5[c] | 31 | 14/45 | 33 | 14/43[b] |
| | | | 9.5[c] | 62 | 28/45 | 65 | 28/43[b] |
| | Selmi *et al*. 1990 (USA) | $HAMD_{17}$ ⩽6[a] | 6 | 8 | 1/12 | 8 | 1/12 |
| | | | 14 | 8 | 1/12 | 8 | 1/12 |
| | | BDI ⩽9 | 6 | 17 | 2/12 | 17 | 2/12 |
| | | | 14 | 8 | 1/12 | 8 | 1/12 |
| | Shaw, 1977 (Canada) | BDI <10[a] | 4 | 0 | 0/8 | 0 | 0/8 |
| | Wierzbicki & Bartlett, 1987 (USA) | BDI ⩽7[a] | 6 | 10 | 2/20 | 10 | 2/20 |
| Older adults | Arean *et al*. 1993 (USA) | Diagnostic criteria no longer met (SADS, DSM-III-R)[a] | 12 | 10 | 2/20 | 10 | 2/20 |
| | Strachowski *et al*. 2008 (USA) | Diagnostic criteria no longer met (DISH, DSM-IV)[a] | 16 | 4 | 1/25 | 4 | 1/25 |
| Children and adolescents | Clarke *et al*. 1999 (USA) | Diagnostic criteria no longer met for ⩾2 weeks (LIFE, DSM-III-R)[a] | 8 | 36 | 13/36 | 48 | 13/27 |
| | Diamond *et al*. 2002 (USA) | Diagnostic criteria no longer met (K-SADS-P, DSM-III-R)[a] | 6 | 44 | 7/16 | 47 | 7/15 |
| | | BDI ⩽9[a] | 6 | 19 | 3/16[d] | 20 | 3/15 |
| | Kowalenko *et al*. 2005 (Australia) | CDI <20[a] | 10 | 18 | 8/44[d] | 20 | 8/41 |
| | Reynolds & Coats, 1986 (USA) | BDI <10[a] | 5 | 0 | 0/10 | 0 | 0/10[e] |
| | | | 10 | 40 | 4/10 | 44 | 4/9 |
| | Weisz *et al*. 1997 (USA) | CDRS-R normal range (⩽1 S.D. above mean for non-clinical school norm)[a] | 8 | 22 | 7/32 | 23 | 7/32 |
| | | | 47[f] | 25 | 8/32 | 24 | 8/32[g] |
| | | CDI normal range (⩽1 S.D. above mean for non-clinical school norm) | 8 | 16 | 5/32 | 16 | 5/32 |
| | | | 47[f] | 31 | 10/32 | 31 | 10/32[g] |
| **Primary-care physician detection studies** | | | | | | | |
| General adults | Goldberg *et al*. 1998 (15 sites, 14 countries) | Diagnostic criteria no longer met (CIDI-PC, DSM-IV) | 52[f] | 31 | 100/323 | 42 | 100/240 |
| | Rost *et al*. 1998 (USA) | ⩽2 of the nine criteria for major depression in the past 2 weeks (DIS) | 26[f] | N.A. | | 16[h] | 5/32 |
| | | | 52[f] | N.A. | | 37[h] | 12/32 |
| | Schulberg *et al*. 1987 (USA) | Diagnostic criteria no longer met (DIS, DSM-III) | 26[f] | N.A. | | 69 | 9/13 |
| | Simon *et al*. 1999 (15 sites, 14 countries) | Diagnostic criteria no longer met (CIDI-PC, ICD-10) | 52[f] | | | | |

https://doi.org/10.1017/S0033291712001717 Published online by Cambridge University Press

**Table 2** (*cont.*)

| Study population | Author, year (country) | Remission definition | Follow-up (weeks) | Proportion remitted | | | |
|---|---|---|---|---|---|---|---|
| | | | | Total sample | | Followed up | |
| | | | | % | *n* | % | *n* |
| | | Baseline mild | | – | | 82[h] | –/–[i] |
| | | Baseline moderate | | – | | 75[h] | –/–[i] |
| | | Baseline severe | | – | | 58[h] | –/–[i] |
| Older adults | Licht-Strunk *et al*. 2009 (The Netherlands) | MADRS <10 and diagnostic criteria no longer met (PRIME-MD, DSM-IV) | 26[f] | N.A. | | 22 | 11/49 |
| | | | 52[f] | N.A. | | 37 | 18/49 |
| **Observational cohort studies in treatment settings** | | | | | | | |
| General adults | Posternak *et al*. 2006 (USA) | Eight consecutive weeks of no or minimal symptoms (LIFE: PSR of 1 or 2, RDC) | 2[j] | N.A. | | 23 | 19/84 |
| | | | 6[j] | N.A. | | 37 | 31/84 |
| | | | 10.5[j] | N.A. | | 52 | 44/84 |
| | | | 19.5[j] | N.A. | | 67 | 56/84 |
| | | | 39[j] | N.A. | | 85 | 71/84 |
| | | | 78[j] | N.A. | | 89 | 75/84 |
| | Schulberg *et al*. 1997 (USA) | HAMD ⩽7 | 35[f] | N.A. | | 20 | 5/25 |
| **Community-based epidemiological cohort studies** | | | | | | | |
| General adults | Wang, 2004 (Canada) | No recurrent MDE at follow-up (CIDI-SFMD, DSM-IV) | 78[j] | N.A. | | 88 | 302/337 |
| | | Low distress group (K6 = 0–7) | | N.A. | | 90 | 149/162 |
| | | Moderate distress group (K6 = 8–12) | | N.A. | | 85 | 95/111 |

BDI, Beck Depression Inventory; CIDI-PC, Composite International Diagnostic Interview (primary-care version); CIDI-SFMD, Composite International Diagnostic Interview – Short Form for Major Depression; CDI, Children's Depression Inventory; CDRS-R, Children's Depression Rating Scale (revised); DIS, Diagnostic Interview Schedule; DISH, Depression Interview and Structured Hamilton; HAMD, Hamilton Rating Scale for Depression (version unspecified); $HAMD_{17}$, 17-item Hamilton Rating Scale for Depression; K6, non-specific psychological distress scale; K-SADS-P, Schedule for Affective Disorders and Schizophrenia for school-age children – present and lifetime version; LIFE, Longitudinal Interval Follow-up Evaluation; MADRS, Montgomery–Asberg Depression Rating Scale; MDE, major depressive episode; N.A., not applicable because untreated status determined at follow-up; PRIME-MD, Primary-Care Evaluation of Mental Disorders; PSR, Psychiatric Status Rating; RDC, Research Diagnostic Criteria; SADS, Schedule for Affective Disorders and Schizophrenia; S.D., standard deviation.

[a] Data corresponding to this remission definition used in primary analysis.

[b] Denominator assumed to be number of participants followed up.

[c] Average of reported follow-up range: 4–7 and 8–11 weeks respectively.

[d] Denominator assumed to be number of participants at baseline.

[e] $n = 1$ not lost to follow-up until 10 weeks.

[f] Converted from time in months, assuming 4.35 weeks in a month.

[g] Denominator assumes no loss to follow-up.

[h] Weighted proportion.

[i] Unable to derive sample size from weighted proportion.

[j] Mid-point between the previous follow-up time-point and that of the reported remission estimate.

https://doi.org/10.1017/S0033291712001717 Published online by Cambridge University Press

Two-thirds (62%) operationalized remission as a below-threshold symptom severity instrument score.

One study (Selmi *et al*. 1990), which included participants with RDC intermittent depressive disorder (17% of the sample) and used a conservative definition of remission (a HAMD score of $\leqslant 6$ rather than the usual $\leqslant 7$), reported a lower proportion remitted than studies of similar duration. Conversely, another (Clarke *et al*. 1999) included participants with DSM-III-R dysthymia (12.5% of the sample) but reported a relatively high proportion remitted. Two studies (Wierzbicki & Bartlett, 1987; Pace & Dixon, 1993) targeted mild-to-moderate major depression by capping the upper bound of study entry symptom severity scores; the former reported a relatively high proportion remitted for studies of similar periods. One study (Wierzbicki & Bartlett, 1987) used a conservative remission threshold (a BDI score of $\leqslant 7$ rather than the usual $<10$) and reported a lower than expected proportion remitted. Studies of older adults also reported lower proportions remitted than those in samples of adults or young people.

Wait-list studies tended to be constrained by convenience samples; failure to document/report treatment rates during follow-up and failure/inability to report diagnostic composition (see online supplementary material for evaluation summary). Patterns of remission were not systematic across wait-list studies that were not evaluated highly; however, two (Shaw, 1977; Reynolds & Coats, 1986) reported no remitted participants after periods of 4 and 5 weeks respectively.

### *Primary-care physician detection studies*

Five samples were drawn from primary-care physician detection studies (see Tables 1 and 2). Sample sizes varied (mean = 125 participants, range 13–240). All reported remission data for follow-up periods of 6 months to 1 year and defined remission using rescinded diagnoses. All samples comprised adults. One study of older adults reported a relatively low proportion remitted.

By design, two primary-care detection studies (Rost *et al*. 1998; Licht-Strunk *et al*. 2009) ensured untreated status by classifying the sample at follow-up. One study (Schulberg *et al*. 1987) reported that 31% of the 'untreated' sample received interventions during follow-up, at least some of which were specific to depression. The proportion remitted was high relative to studies of similar duration. Another two studies (Goldberg *et al*. 1998; Simon *et al*. 1999) classified treatment status at baseline, without describing the proportion who may have received treatment during follow-up. Nonetheless, detection study methodology was generally evaluated positively (see online supplementary material).

### *Observational cohort studies*

Three observational cohort studies were identified (see Tables 1 and 2). The first (Posternak *et al*. 2006) defined remission using a rescinded diagnosis and reported proportions remitting after periods ranging between 4 weeks and 2 years. Although recruitment took place in academic medical settings, the intake episode was not analyzed and participants were regarded as 'non-treatment seekers' by choosing not to obtain somatic therapy. Remission rates for this study were high relative to estimates from studies with comparable follow-up durations. Another study (Schulberg *et al*. 1997) was the prospective component of an RCT that recruited participants from primary-care settings. Remission was defined as scores below the conventional HAMD threshold of $\leqslant 7$ after 8 months. This follow-up period was shorter than other studies of similar duration (1 year) and the remission rate was comparatively lower. The final study (Wang, 2004) was the only longitudinal, community-based epidemiological study identified. Rescinded diagnoses defined remission after a 2-year period but service utilization was measured only in the preceding year.

These studies were of relatively good methodological quality. The most common methodological problem was the failure to report the extent of loss to follow-up (see online supplementary material).

### *Predicting the probability of remission from untreated depression*

A regression model for predicting remission from untreated depression was developed through an iterative process of model building and testing. Two studies (Goldberg *et al*. 1998; Simon *et al*. 1999) reported data from overlapping samples. Only the data reported by Goldberg *et al*. (1998) were entered into the model as these authors reported on the entire sample, unstratified by severity. Of several studies found to have large residual values (Pace & Dixon, 1993; Weisz *et al*. 1997; Goldberg *et al*. 1998; Wang, 2004; Posternak *et al*. 2006), three (Weisz *et al*. 1997; Goldberg *et al*. 1998; Wang, 2004) were identified as having greater than expected influence over the model, with one (Wang, 2004) exerting substantially greater influence than others, as judged by Cook's distance. Review of the methodological aspects of these studies showed that the latter study involved the longest follow-up duration (78 weeks) and was the only community-based epidemiological observational cohort study. No obvious methodological differences were found in the

https://doi.org/10.1017/S0033291712001717 Published online by Cambridge University Press

**Table 3.** *Odds ratios (ORs) for remission in untreated depression*

| Covariate | No. of studies | No. of cases | OR | Robust S.E. | *z* | $p>z$ | 95% CI |
|---|---|---|---|---|---|---|---|
| Duration of follow-up (weeks) | 19 | 749 | 0.97 | 0.01 | −3.05 | <0.01 | 0.95–0.99 |
| Age group | | | | | | | |
| Adults | 11 | 531 | 1.00 | | | | |
| Child/adolescent | 5 | 124 | 1.88 | 0.52 | 2.25 | 0.02 | 1.09–3.25 |
| Older adults | 3 | 94 | 0.52 | 0.14 | −2.42 | 0.02 | 0.31–0.88 |
| Type of sample | | | | | | | |
| Primary-care | 6 | 443 | 1.00 | | | | |
| Wait-listed | 13 | 306 | 0.09 | 0.04 | −5.08 | <0.01 | 0.03–0.23 |

S.E., Standard error; CI, confidence interval.

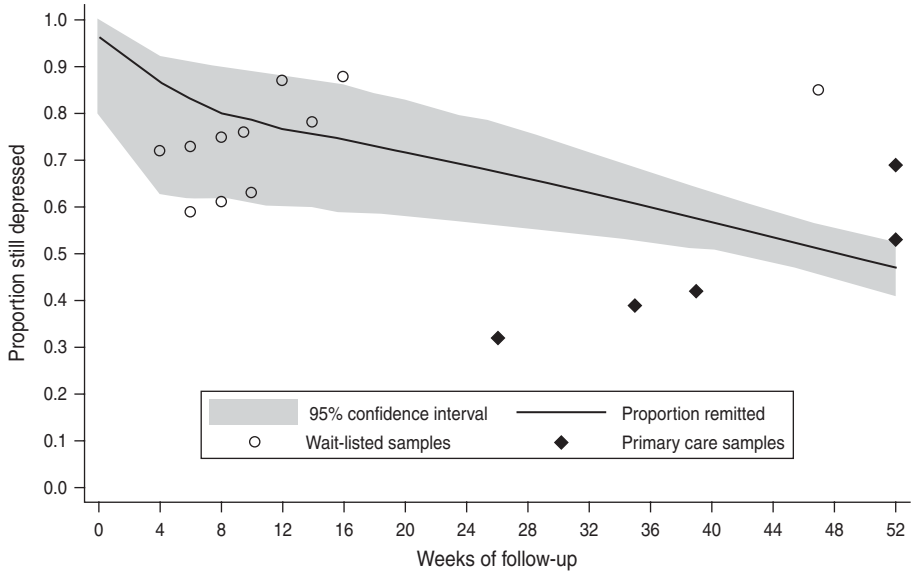


**Fig. 2.** Adjusted proportion of cases still depressed, by time, with 95% confidence intervals and weighted average proportions.

other studies that would justify their exclusion. Therefore, subsequent analyses excluded the Wang (2004) study.

The final multivariate regression model is shown in Table 3. The factor most strongly associated with remission was study type. After controlling for other variables, the probability of remission at 12 months was significantly lower for wait-listed samples compared to samples ascertained through primary-care settings. Study follow-up duration was also significantly associated with the probability of remission such that, for each additional week of follow-up, the probability of remitting in that week diminished by 3%. The probability of remission was significantly higher in child and adolescent groups and lower in older adults compared to general adult samples, although analyses for these age groups were based on few studies ($n=5$ and $n=3$ respectively). No significant association was found between probability of remission and gender composition or type of diagnostic measure.

To investigate remission of adult cases of depression that are representative of the average community case, the predicted remission was adjusted to these reference categories. Figure 2 shows the adjusted distribution of the proportion of cases still depressed by time over a 12-month period for adults recruited from primary-care studies, based on the regression model. For every 100 adults with depression, 77 will still be depressed at 3 months, 68 at 6 months and 47 at 1 year. This suggests that, without treatment, 23% adult cases of depression will remit within 3 months, 32% within 6 months and 53% within 1 year.

https://doi.org/10.1017/S0033291712001717 Published online by Cambridge University Press

A sensitivity analysis tested the regression model with and without the influential study (Wang, 2004). The regression equations differed considerably. As the model excluding the Wang study produced a lower Akaike Information Criterion (AIC) value (948 *v*. 1260), this model was considered more accurate.

### *Remission by severity*

Two studies stratified remission proportions by disorder severity but methodological differences precluded pooling their data. Simon *et al*. (1999) applied ICD-10 depressive episode severity criteria (mild, moderate, severe) and followed up participants after 1 year. Wang (2004) used thresholds on a non-specific psychological distress scale (K6), corresponding to low and moderate distress, and followed up participants after 2 years. Nonetheless, qualitatively similar results were observed. The proportions remitted decreased as depression severity increased. An estimated 82% (Simon *et al*. 1999) and 90% (Wang, 2004) were in remission from mild depression after 1- and 2-year follow-ups respectively. This decreased to 75% (Simon *et al*. 1999) and 85% (Wang, 2004) in remission from moderate depression after 1- and 2-year follow-ups respectively. Simon *et al*. (1999) estimated remission from untreated severe depression to be 58% after 1 year.

## Discussion

Our analysis, based on 19 studies, suggests that approximately 53% of prevalent cases of untreated major depression will remit spontaneously in a given year. This is almost half that of estimates (>90%) from community-based epidemiological studies that include both treated and untreated cases (McLeod *et al*. 1992; Lewinsohn *et al*. 1994; Kendler *et al*. 1997). There is also evidence that remission rates may depend on disorder severity, with rates among people with severe disorders being 20–30% lower than mild-to-moderate disorders. However, this observation is based on only two studies and should be interpreted with caution.

Can this information inform population treatment targets for depression? In the first instance they suggest that even if it were possible to achieve total coverage by providing sufficient services, this is not an appropriate target because of high natural remission rates. However, remission data alone cannot be used to set a treatment target. Duration of remission, personal preferences, available resources and financial and other treatment barriers may also play a part, but were beyond the scope of this paper to investigate.

In Australia, the National Surveys of Mental Health and Wellbeing estimated the treatment rate (defined as consulting a health professional) for affective disorders to be 66.8% in 1997 (Andrews *et al*. 2001) and 58.6% in 2007 (Burgess *et al*. 2009). Many of those who remained untreated said they did not require interventions. The reasons for this are complex (Meadows & Burgess, 2009) but support the view that prevalence alone is not an adequate indicator of need for treatment. An equity-based approach promotes the view that treatment rates for depression should match those of other common physical disorders that cause comparable levels of disability. In Australia, this would suggest rates of approximately 90% if disorders such as rheumatoid arthritis and osteoarthritis (treatment rate, 93%; Andrews *et al*. 2006) and asthma (89%; Simonella *et al*. 2006) were chosen. However, setting rates for face-to-face treatment at 90% would require considerable resource allocation out of reach for most countries. Fortunately, the increasing evidence base supporting online therapies (Griffiths *et al*. 2010) suggests that increased treatment coverage might be achieved, at least in part, with judicious use of inexpensive e-therapies.

The findings of this study are also relevant to clinicians. Spontaneous recovery from depression is high, and attribution of recovery to the effects of treatment is likely to be overestimated. A short period without active treatment in mild-to-moderate depression would seem defensible, given the analysis was based largely on samples of consenting out-patient and wait-listed groups. This is consistent with the stepped-care model endorsed by the UK National Institute for Health and Clinical Excellence (NICE) guidelines (NICE, 2009), which recommend a watchful waiting period for mild depression, depending on individual case details. However, our findings provide little information for clinicians regarding relapse prevention, given that the scope of the paper was limited to time to remission.

### *Threats to validity*

To our knowledge this is the most comprehensive review of remission from untreated depression conducted since 2001. The review conformed with PRISMA guidelines (Moher *et al*. 2009) and directed considerable effort towards review-level bias prevention. The search strategy did not impose language limitations, captured a variety of informative study designs and enabled unpublished findings to be included. The population of interest and acceptable criteria for defining remission were specified *a priori*. A purpose-designed methodological assessment

https://doi.org/10.1017/S0033291712001717 Published online by Cambridge University Press

template allowed us to describe potential sources of variation in detail.

Nonetheless, potential sources of bias should be considered. First, we could not be certain that all participants remained untreated for the duration of follow-up. For example, wait-list control groups may have received interventions for depression outside the study protocol; however, such data were rarely reported. Only one wait-list study (Clarke *et al*. 1999) measured other treatments received and excluded those participants from their analysis. In three of the five primary-care detection studies (Schulberg *et al*. 1987; Goldberg *et al*. 1998; Simon *et al*. 1999), participants were classified as undetected/untreated at baseline and subsequent treatment was not reported by two of these (Goldberg *et al*. 1998; Simon *et al*. 1999). Studies also differed in the breadth of treatments considered to be restricted during follow-up; for example, somatic treatments only or somatic and psychological treatments. Potentially including treated individuals may have resulted in an overestimate of the remission rate.

Second, self-selecting samples recruited in treatment settings may have affected the generalizability of results. A perceived need for care has been associated with greater symptom severity (Sareen *et al*. 2005; van Beljouw *et al*. 2010*a*,*b*), which may have biased these samples towards lower spontaneous remission rates. On the contrary, more severe cases may have been excluded. Most wait-list studies excluded those describing suicidal ideation and/or requiring a higher level of care, and two (Wierzbicki & Bartlett, 1987; Pace & Dixon, 1993) specifically targeted mild-to-moderate major depression. Several detection studies (Goldberg *et al*. 1998; Simon *et al*. 1999; Licht-Strunk *et al*. 2009) showed that physicians were more likely to recognize and treat depression of greater severity, thus these individuals were less likely to appear in the 'untreated' group. On balance, the preponderance of wait-list and primary-care samples probably resulted in an over-representation of people with less severe depression and may have biased remission estimates in a positive direction.

Similarly, the inclusion of studies using symptom scales to determine diagnosis potentially affected the generalizability of our results. To maximize sample size and coverage of the targeted research designs, we included studies that derived diagnoses from scales with appropriate cross-walks to diagnostic systems. However, the threshold for diagnosis using symptom scales is not absolute. Symptom scales may admit more people to the diagnosis-positive group, particularly those with mild disorders. Assuming mild depression is more likely to remit, this may in turn overestimate remission rates. However, we did not find a statistically significant association between remission and type of diagnostic instrument (symptom scales *versus* diagnostic interview).

Third, some studies may have included a minority of participants with chronic mood disorders, which may have negatively biased remission estimates. However, diagnostic breakdown was reported in only three studies (Selmi *et al*. 1990; Rost *et al*. 1998; Clarke *et al*. 1999).

Fourth, loss to follow-up potentially influenced our results. Studies have shown greater social impairment and symptom duration to be positively (Sonawalla *et al*. 2002) and negatively (Simon & Ludman, 2010) associated with early drop-out from treatment. However, little is known about factors associated with drop-out from wait-listed groups due to inadequate reporting practices (Cisler *et al*. 2007). As such, it is difficult to speculate on the effects of attrition on the present findings.

Fifth, more comprehensive definitions of remission require improvements in psychosocial functioning and physiological factors in addition to symptom reduction (Keller, 2003; Israel, 2006; Zimmerman *et al*. 2006, 2008). The consistent application of a comprehensive definition of remission across studies would have been ideal. However, the identified studies operationalized remission using information from instruments focused on symptom levels. These may have overestimated remission rates because remission criteria did not incorporate psychosocial functioning and physiological factors.

Finally, we noted the lack of spontaneous remission data available from epidemiological studies with a longitudinal component. These studies can arguably provide the most accurate estimate of remission as they include individuals who have not presented to health services, and who may or may not have a perceived need for care. It is possible that other researchers have access to such data, but they have not been reported. However, we did not anticipate publication bias to be a problem for this review because the untreated samples were generally not the focus of interest in the studies sought.

## Conclusions

The findings from this suggest that, although treatment is likely to facilitate earlier remission, just over half of those with a major depressive episode will remit within a year without intervention. The high rate of spontaneous remission observed among wait-listed and primary-care samples in this review has implications for service planners and clinicians. For service planners it is undesirable to expect 100% treatment coverage because many people with less severe major

https://doi.org/10.1017/S0033291712001717 Published online by Cambridge University Press

depression will remit before access to services is feasible. For clinicians, a short period without active treatment for people with mild-to-moderate depression (NICE, 2009) seems defensible. Resources should be directed towards those with greatest need, for example those experiencing more severe depression and those whose symptoms are likely to persist or reoccur. Remission from untreated depression remains an overlooked area of research. There is a need for natural remission rates from high-quality population-level studies to be examined. Analyses of data from prospective cohort studies could focus on elucidating factors that may determine the likelihood of remission from depression among untreated individuals.

## Supplementary material

For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S0033291712001717.

## Acknowledgments

We thank G. Williams and A. Higginson from the University of Queensland for statistical advice. We also thank M. Dewey for the detailed statistical report he provided as a previous reviewer.

## Declaration of Interest

None.

## References

**Allen JJ, Schnyer RN, Hitt SK** (1998). The efficacy of acupuncture in the treatment of major depression in women. *Psychological Science* **9**, 397–401.

**Allen JJB, Schnyer RN, Chambers AS, Hitt SK, Moreno FA, Manber R** (2006). Acupuncture for depression: a randomized controlled trial. *Journal of Clinical Psychiatry* **67**, 1665–1673.

**Alonso J, Angermeyer M, Bernert S, Bruffaerts R, Brugha T, Bryson H, de Girolamo G, de Graaf R, Demyttenaere K, Gasquet I, Haro J, Katz S, Kessler R, Kovess V, Lepine J, Ormel J, Polidori G, Russo L, Vilagurt G** (2004). Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatrica Scandinavica* **109**, 21–27.

**Andrews G** (2007). *Tolkien II: A Needs-based, Costed, Stepped-care Model for Mental Health Services*. World Health Organization, Collaborating Centre for Classification in Mental Health: Sydney.

**Andrews G, Issakidis C, Carter G** (2001). Shortfall in mental health service utilisation. *British Journal of Psychiatry* **179**, 417–425.

**Andrews G, Simonella L, Lapsley H, Sanderson K, March L** (2006). Evidence-based medicine is affordable: the cost-effectiveness of current compared with optimal treatment in rheumatoid and osteoarthritis. *Journal of Rheumatology* **33**, 671–680.

**Arean PA, Perri MG, Nezu AM, Schein RL, Christopher F, Joseph TX** (1993). Comparative effectiveness of social problem-solving therapy and reminiscence therapy as treatments for depression in older adults. *Journal of Consulting and Clinical Psychology* **61**, 1003–1010.

**Azorin JM** (1995). Spontaneous course of depression. *Encephale* **21**, 31–34.

**Beck AT, Steer RA** (1988). Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clinical Psychology Review* **8**, 77–100.

**Buis ML** (2007). Predict and adjust with logistic regression. *Stata Journal* **7**, 221–226.

**Burgess PM, Pirkis JE, Slade TN, Johnston AK, Meadows GN, Gunn JM** (2009). Service use for mental health problems: findings from the 2007 National Survey of Mental Health and Wellbeing. *Australian and New Zealand Journal of Psychiatry* **43**, 615–623.

**Calabria B, Degenhardt L, Briegleb C, Vos T, Hall W, Lynskey M, Callaghan B, Rana U, McLaren J** (2010). Systematic review of prospective studies investigating 'remission' from amphetamine, cannabis, cocaine or opioid dependence. *Addictive Behaviors* **35**, 741–749.

**Cisler JM, Barnes AC, Farnsworth D, Sifers SK** (2007). Reporting practices of dropouts in psychological research using a wait-list control: current state and suggestions for improvement. *International Journal of Methods in Psychiatric Research* **16**, 34–42.

**Clarke GN, Rohde P, Lewinsohn PM, Hops H, Seeley JR** (1999). Cognitive-behavioral treatment of adolescent depression: efficacy of acute group treatment and booster sessions. *Journal of the American Academy of Child and Adolescent Psychiatry* **38**, 272–279.

**Coryell W, Endicott J, Winokur G, Akiskal H, Solomon D, Leon A, Mueller T, Shea T** (1995). Characteristics and significance of untreated major depressive disorder. *American Journal of Psychiatry* **152**, 1124–1129.

**Diamond GS, Reis BF, Diamond GM, Siqueland L, Isaacs L** (2002). Attachment-based family therapy for depressed adolescents: a treatment development study. *Journal of the American Academy of Child and Adolescent Psychiatry* **41**, 1190–1196.

**Fox HA** (2002). The natural course of depression: Kraeplin and beyond. *Harvard Review of Psychiatry* **10**, 249–253.

**Frank E, Prien RF, Jarrett RB, Keller MB, Kupfer DJ, Lavori PW, Rush AJ, Weissman MM** (1991). Conceptualization and rationale for consensus definitions of terms in major depressive disorder: remission, recovery, relapse, and recurrence. *Archives of General Psychiatry* **48**, 851–855.

**Goldberg D, Privett M, Ustun B, Simon G, Linden M** (1998). The effects of detection and treatment on the outcome of major depression in primary care: a naturalistic study in 15 cities. *British Journal of General Practice* **48**, 1840–1844.

**Griffiths K, Farrer L, Christensen H** (2010). The efficacy of internet interventions for depression and anxiety disorders: a review of randomised controlled trials. *Medical Journal of Australia* **192**, S4–S11.

**Grilo CM, Sanislow CA, Shea MT, Skodol AE, Stout RL, Gunderson JG, Yen S** (2005). Two-year prospective naturalistic study of remission from major depressive disorder as a function of personality disorder comorbidity. *Journal of Consulting and Clinical Psychology* **73**, 78–85.

**Hohman LB** (1938). A review of 144 cases of affective disorders after seven years. *American Journal of Psychiatry* **94**, 304–308.

**Israel JA** (2006). Remission in depression: definition and initial treatment approaches. *Journal of Psychopharmacology* **20**, 5–10.

**Keller MB** (2003). Past, present, and future directions for defining optimal treatment outcome in depression: remission and beyond. *Journal of the American Medical Association* **289**, 3152–3160.

**Kendler KS, Walters EE, Kessler RC** (1997). The prediction of length of major depressive episodes: results from an epidemiological sample of female twins. *Psychological Medicine* **27**, 107–117.

**Kessler R, Chiu W, Demler O, Walters E** (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry* **62**, 617–627.

**Kowalenko N, Rapee RM, Simmons J, Wignall A, Hoge R, Whitefield K, Starling J, Stonehouse R, Baillie AJ** (2005). Short-term effectiveness of a school-based early intervention program for adolescent depression. *Clinical Child Psychology and Psychiatry* **10**, 493–507.

**Krøgsboll LT, Hróbjartsson A, Gøtzsche PC** (2009). Spontaneous improvement in randomised clinical trials: meta-analysis of three-armed trials comparing no treatment, placebo and active intervention. *BMC Medical Research Methodology* **9**, 1.

**Lewinsohn PM, Clarke GN, Seeley JR, Rohde P** (1994). Major depression in community adolescents: age at onset, episode duration, and time to recurrence. *Journal of the American Academy of Child and Adolescent Psychiatry* **33**, 809–818.

**Licht-Strunk E, Beekman ATF, de Haan M, van Marwijk HWJ** (2009). The prognosis of undetected depression in older general practice patients. A one year follow-up study. *Journal of Affective Disorders* **114**, 310–315.

**Lim D, Sanderson K, Andrews G** (2000). Lost productivity among full-time workers with mental disorders. *Journal of Mental Health Policy and Economics* **3**, 139–146.

**McGrath J, Saha S, Welham J, El Saadi O, MacCauley C, Chant D** (2004). A systematic review of the incidence of schizophrenia: the distribution of rates and the influence of sex, urbanicity, migrant status and methodology. *BMC Medicine* **2**, 13.

**McLeod JD, Kessler RC, Landis KR** (1992). Speed of recovery from major depressive episodes in a community sample of married men and women. *Journal of Abnormal Psychology* **101**, 277–286.

**Meadows G, Burgess P** (2009). Perceived need for mental health care: findings from the 2007 Australian Survey of Mental Health and Wellbeing. *Australian and New Zealand Journal of Psychiatry* **43**, 624–634.

**Mendlewicz J** (2009). Defining remission in depression: the challenge of complete recovery. *Medicographia* **31**, 113–117.

**Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group** (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine* **6**, e1000097.

**Murray CJ, Lopez AD** (1997). Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. *Lancet* **349**, 1436–1442.

**Musial F, Klosterhalfen S, Enck P** (2007). Placebo responses in patients with gastrointestinal disorders. *World Journal of Gastroenterology* **13**, 3425–3429.

**NICE** (2009). *NICE Clinical Guideline 90. Depression: The Treatment and Management of Depression in Adults* (*partial update of NICE Clinical Guideline 23*). The British Psychological Society and the Royal College of Psychiatrists: London, UK.

**Ormel J, Oldehinkel T, Brilman E, vanden Brink W** (1993). Outcome of depression and anxiety in primary care. A three-wave 3 1/2-year study of psychopathology and disability. *Archives of General Psychiatry* **50**, 759–766.

**Pace TM, Dixon DN** (1993). Changes in depressive self-schemata and depressive symptoms following cognitive therapy. *Journal of Counseling Psychology* **40**, 288–294.

**Posternak MA, Miller I** (2001). Untreated short-term course of major depression: a meta-analysis of outcomes from studies using wait-list control groups. *Journal of Affective Disorders* **66**, 139–146.

**Posternak MA, Solomon DA, Leon AC, Mueller TI, Shea M, Endicott J, Keller MB** (2006). The naturalistic course of unipolar major depression in the absence of somatic therapy. *Journal of Nervous and Mental Disease* **194**, 324–329.

**Posternak MA, Zimmerman M** (2000). Short-term spontaneous improvement rates in depressed outpatients. *Journal of Nervous and Mental Disease* **188**, 799–804.

**Reynolds WM, Coats KI** (1986). A comparison of cognitive-behavioral therapy and relaxation training for treatment of depression in adolescents. *Journal of Consulting and Clinical Psychology* **54**, 653–660.

**Rost K, Zhang M, Fortney J, Smith J, Coyne J, Smith G** (1998). Persistently poor outcomes of undetected major depression in primary care. *General Hospital Psychiatry* **20**, 12–20.

**Sareen J, Cox B, Afifi T, Clara I, Yu B** (2005). Perceived need for mental health treatment in a nationally representative Canadian sample. *Canadian Journal of Psychiatry* **50**, 643–651.

**Schorer C** (1970). Improvement with and without psychotherapy. *Diseases of the Nervous System* **31**, 155–160.

**Schorer C, Lowinger P, Sullivan T, Hartlaub GH** (1968). Improvement without treatment. *Diseases of the Nervous System* **29**, 100–104.

**Schulberg HC, Block MR, Madonia MJ, Scott PC, Lave JR, Rodriguez E, Coulehan JL** (1997). The 'usual care' of major depression in primary care practice. *Archives of Family Medicine* **6**, 334–339.

**Schulberg HC, McClelland M, Gooding W** (1987). Six-month outcomes for medical patients with major

depressive disorders. *Journal of General Internal Medicine* **2**, 312–317.

**Selmi PM, Klein MH, Greist JH, Sorrell SP, Erdman HP** (1990). Computer-administered cognitive-behavioral therapy for depression. *American Journal of Psychiatry* **147**, 51–56.

**Shaw BF** (1977). Comparison of cognitive therapy and behavior therapy in the treatment of depression. *Journal of Consulting and Clinical Psychology* **45**, 543–551.

**Simon GE, Fleck M, Lucas R, Bushnell DM** (2004). Prevalence and predictors of depression treatment in an international primary care study. *American Journal of Psychiatry* **161**, 1626–1634.

**Simon GE, Goldberg D, Tiemens BG, Ustun TB** (1999). Outcomes of recognized and unrecognized depression in an international primary care study. *General Hospital Psychiatry* **21**, 97–105.

**Simon GE, Ludman EJ** (2010). Predictors of early dropout from psychotherapy for depression in community practice. *Psychiatric Services* **61**, 684–689.

**Simonella L, Marks G, Sanderson K, Andrews G** (2006). Cost-effectiveness of current and optimal treatment for adult asthma. *Internal Medicine Journal* **36**, 244–250.

**Slade T, Johnston A, Teesson M, Whiteford H, Burgess P, Pirkis J, Saw S** (2009). *The Mental Health of Australians 2: Report on the 2007 National Survey of Mental Health and Wellbeing*. Commonwealth Department of Health and Ageing: Canberra.

**Sonawalla SB, Farabaugh AH, Leslie VM, Pava JA, Matthews JD, Fava M** (2002). Early drop-outs, late drop-outs and completers: differences in the continuation phase of a clinical trial. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **26**, 1415–1419.

**Strachowski D, Khaylis A, Conrad A, Neri E, Spiegel D, Taylor CB** (2008). The effects of cognitive behavior therapy on depression in older patients with cardiovascular risk. *Depression and Anxiety* **25**, E1–E10.

**van Beljouw I, Verhaak P, Cuijpers P, van Marwijk H, Penninx B** (2010*a*). The course of untreated anxiety and depression and the determinants of poor one-year outcome: a one-year cohort study. *BMC Psychiatry* **10**, 86.

**van Beljouw I, Verhaak P, Prins M, Cuijpers P, Penninx B, Bensing J** (2010*b*). Reasons and determinants for not receiving treatment for common mental disorders. *Psychiatric Services* **61**, 250–257.

**Wang J** (2004). A longitudinal population-based study of treated and untreated major depression. *Medical Care* **42**, 543–550.

**Weisz JR, Thurber CA, Sweeney L, Proffitt VD, LeGagnoux GL** (1997). Brief treatment of mild-to-moderate child depression using primary and secondary control enhancement training. *Journal of Consulting and Clinical Psychology* **65**, 703–707.

**WHO** (2008). *The Global Burden of Disease 2004 Update*. World Health Organization: Geneva.

**Wierzbicki M, Bartlett TS** (1987). The efficacy of group and individual cognitive therapy for mild depression. *Cognitive Therapy and Research* **11**, 337–342.

**Zimmerman M, McGlinchey JB, Posternak M, Friedman M, Boerescu D, Attiullah N** (2006). Discordance between self-reported symptom severity and psychosocial functioning ratings in depressed outpatients: implications for how remission from depression should be defined. *Psychiatry Research* **141**, 185–191.

**Zimmerman M, McGlinchey JB, Posternak M, Friedman M, Boerescu D, Attiullah N** (2008). Remission in depressed outpatients: more than just symptom resolution? *Journal of Psychiatric Research* **42**, 797–801.

https://doi.org/10.1017/S0033291712001717 Published online by Cambridge University Press