

# **MANGA STUDY**

**COMPARATIVE EFFICACY AND ACCEPTABILITY  
OF 12 NEW GENERATION ANTIDEPRESSANTS:  
A MULTIPLE TREATMENT META-ANALYSIS**

Andrea Cipriani, Toshiaki A. Furukawa, Georgia Salanti, John R. Geddes, Rachel Churchill,  
Julian Higgins, Hugh McGuire, Norio Watanabe, Alessandra Signoretti,  
Atsuo Nakagawa, Michela Nosè, Ichiro M. Omori, Antonio Veronese and Corrado Barbui

**PROTOCOL**

FINAL VERSION – DECEMBER 2007

## BACKGROUND

Major depressive disorder is the most prevalent psychiatric disease in the general population, affecting more than 16% of adults during their lifetime (Kessler et al., 2003). In 2000 the economic burden of depressive disorders in the US was estimated to be around 80 billion dollars, with more than 30% of these costs being attributable to direct medical expenses (Greenberg et al., 2003). First-generation antidepressants, such as tricyclics, are no longer agents of choice and their place has been taken by second generation agents including selective serotonin reuptake inhibitors (SSRIs), serotonin and norepinephrine reuptake inhibitors (SNRIs), and others (such as mirtazapine, reboxetine and bupropion) (IMS Health, 2004).. Current clinical practice guidelines on the treatment of depressive disorder recommend that a second-generation antidepressant (usually an SSRI) should be the first-line option when drug therapy is indicated for a depressive episode (NICE, 2004).

It is noteworthy that some systematic reviews have found that certain second-generation antidepressants are more efficacious than other drugs both within and between classes (Puech et al., 1997; Smith et al., 2002; Hansen et al., 2005; Cipriani et al., 2006; Papakostas et al., 2007; Watanabe et al., in press). However, these differences are inconsistent across different systematic reviews. A systematic review conducted by the RTI International-University of North Carolina Evidence-based Practice Centre and the Agency for Healthcare Research and Quality (AHRQ) summarized the available evidence on the comparative efficacy, effectiveness, and harms of 12 second-generation antidepressants and conducted meta-analyses for four direct drug-drug comparisons and 62 indirect comparisons between drugs (Gartlehner et al., 2007). Neither direct or indirect comparisons found substantial differences in efficacy among second-generation antidepressants.

However, the main limitation of this review is that authors synthesized the literature qualitatively, augmenting findings with quantitative analyses only if head-to-head data were sufficient. By contrast, indirect evidence can be used not only *in lieu* of direct evidence, but also to supplement it (Song et al., 2003). Moreover, Gartlehner et al limited themselves to English language literature and consequently included only a subset of relevant randomised controlled trials (RCTs).

Multiple treatment meta-analysis (MTM) is a statistical technique that allows both direct and indirect comparisons to be undertaken, even when two of the treatments have not been directly compared (Higgins et al., 1996; Hasselblad et al., 1998; Lumley, 2002). In other words, it is a generalisation of

standard pair-wise meta-analysis for A vs B trials, to data structures that include, for example, A vs B, B vs C, and A vs C trials.

MTM (also known as *network meta-analysis*) can summarise RCTs of several different treatments providing point estimates (together with 95% confidence intervals [CIs]) for their association with a given endpoint, as well as an estimate of incoherence (that is, a measure of how well the entire network fits together, with small values suggesting better internal agreement of the model). MTM has already been used successfully in other fields of medicine (Psaty et al., 2003; Elliott et al., 2007) and two fruitful roles for MTM have been identified (Lu & Ades, 2004):

- (i) to strengthen inferences concerning the relative efficacy of two treatments, by including both direct and indirect comparisons to increase precision and combine both direct and indirect evidence (Salanti et al., in press);
- (ii) to facilitate simultaneous inference regarding all treatments in order for example to select the best treatment. Considering how important comparative efficacy could be for clinical practice and policy making, it is useful to use all the available evidence to estimate potential differences in efficacy among treatments.

The present review will be the main publication of the Meta-Analyses of New Generation Antidepressants (MANGA) project in which a group of researchers within the Cochrane Collaboration Depression, Anxiety and Neurosis Group agreed to systematically review all available evidence for 12 new generation antidepressants, in order to inform clinical practice and mental health policies. We will use MTM to estimate the comparative efficacy and acceptability of the twelve most frequently used new generation antidepressants for acute treatment of major depression.

## **OBJECTIVES**

To compare individual newer generation antidepressants in terms of:

- (1) Response to antidepressant treatment, defined as the proportion of patients who showed 50% or greater reduction in depression severity at 8 weeks.
- (2) Acceptability of treatment, defined as the proportion of patients who left the study early by any cause during the first 8 weeks of treatment.

## **METHODS**

## ***Criteria for considering studies for this review***

### *Types of studies*

RCTs comparing one drug with another (head-to-head studies) within the same group of 12 second-generation antidepressants (namely, bupropion, citalopram, duloxetine, escitalopram, fluoxetine, fluvoxamine, milnacipran, mirtazapine, paroxetine, reboxetine, sertraline, and venlafaxine) as monotherapy in the acute phase treatment of depression will be included. We will include only head-to-head active comparisons, excluding placebo arms where present. Trials in which antidepressants were used as an augmentation strategy will be excluded. Quasi-randomized trials (such as those allocating by using alternate days of the week) will be excluded. For trials which have a crossover design only results from the first randomisation period will be considered.

### *Types of participants*

Patients aged 18 or older, of both sexes with a primary diagnosis of depression. Studies adopting any standardised criteria to define patients suffering from unipolar major depression will be included. Most recent studies are likely to have used DSM-IV (APA 1994) or ICD-10 (WHO 1992) criteria. Older studies may have used ICD-9 (WHO 1978), DSM-III (APA 1980)/DSM-III-R (APA 1987) or other diagnostic systems. ICD-9 is not operationalised, because it has only disease names and no diagnostic criteria, so studies using ICD-9 will be excluded. On the other hand, studies using Feighner criteria or Research Diagnostic Criteria will be included. Studies in which less than 20% of the participants may be suffering from bipolar depression will be included. A concurrent secondary diagnosis of another psychiatric disorder will not be considered as exclusion criteria. Trials in which all participants have a concurrent primary diagnosis of Axis I or II disorders will be excluded. Studies in which all participants have a diagnosis of resistant depression will be excluded. Antidepressant trials in depressive patients with a serious concomitant medical illness will be excluded. RCTs of women with post-partum depression will be also excluded, because post-partum depression appears to be clinically different from major depression (Cooper & Murray, 1998).

## ***Outcome measures***

### *(1) Response to antidepressant treatment*

Response is defined as the proportion of patients who show at 8 weeks a reduction of at least 50% on Hamilton Depression Rating Scale (HDRS) (Hamilton, 1960) or Montgomery-Åsberg Depression Rating Scale (MADRS) (Montgomery, 1979) or who will score 'much improved' or

'very much improved' at the or Clinical Global Impression (CGI) (Guy, 1970), out of the total number of patients randomly assigned to each antidepressant. When all the scores are provided, we will prefer the former measurement for judging response. Furukawa and colleagues have reported the possibility of underreporting the measured outcomes (reporting bias), therefore we will not employ the original author's definitions of response outcomes (Furukawa et al., 2007).

## (2) *Acceptability of treatment*

Treatment discontinuation (acceptability) is defined as the proportion of patients who leave the study early for any reason during the first 8 weeks of treatment, out of the total number of patients randomly assigned to each antidepressant.

## *Search strategy*

All published and unpublished randomized controlled trials that compared the efficacy and acceptability (dropout rate) of one second generation antidepressants with another (see the list of included antidepressants here above) in the treatment of major depression will be identified by searches of the Cochrane Collaboration Depression, Anxiety & Neurosis Review Group Controlled Trials Registers (CCDANDTR-Studies and CCDANCTR-References. This register is compiled from systematic and regularly updated searches of Cochrane Collaboration CENTRAL register, AMED, CINAHL, EMBASE, LiLACS, MEDLINE, UK National Research Register, PSYCINFO, PSYINDEX supplemented with hand searching of 12 conference proceedings (for details, see specialized register section of CCDAN module on Cochrane Library <http://www.mrw.interscience.wiley.com/cochrane/clabout/articles/DEPRESSN/frame.html>).

Trial databases of the following drug-approving agencies - (the Food and Drug Administration (FDA) in the USA, the Medicines and Healthcare products Regulatory Agency (MHRA) in the UK, the European Medicines Agency (EMA) in the EU, the Pharmaceuticals and Medical Devices Agency (PMDA) in Japan, the Therapeutic Goods Administration (TGA) in Australia) and ongoing trial registers (clinicaltrials.gov in the USA, ISRCTN and National Research Register in the UK, Netherlands Trial Register in the Netherlands, EUDRACT in the EU, UMIN-CTR in Japan and the Australian Clinical Trials Registry in Australia) will be hand-searched for published, unpublished and ongoing controlled trials.

No language restrictions will be applied. The following phrase will be used: [*depress\** or *dysthymi\** or *adjustment disorder\** or *mood disorder\** or *affective disorder* or *affective symptoms*] and combined with a list of 12 specific second-generation antidepressants (bupropion, citalopram, duloxetine, escitalopram, fluoxetine, fluvoxamine, milnacipran, mirtazapine, paroxetine, reboxetine, sertraline, and

venlafaxine). All relevant authors will be contacted to supplement the incomplete report of the original papers.

We are aware that there are many trials carried out in China (Chakrabarti et al., 2007). However, for many of these studies only incomplete or conflicting information is available. In an effort to avoid the potential biases that may be introduced by including these trials without further information, we will list them as “awaiting assessment” for transparency.

### ***Study selection and data extraction***

Two persons will independently review references and abstracts retrieved by the search. If both reviewers agree that the trial doesn't meet eligibility criteria, we will exclude it. We will obtain the full text of all remaining articles and use the same eligibility criteria to determine which, if any, to exclude at this stage. Any disagreements will be solved via discussion with a third member of the reviewing team.

Two reviewers will then independently read each article, evaluate the completeness of the data abstraction, and confirm the quality rating. We will design and use a structured data abstraction form to ensure consistency of appraisal for each study. Information extracted will include study characteristics (such as lead author, publication year, journal), participant characteristics (such as diagnostic criteria for depression, age range, setting, diagnosis of bipolar depression), intervention details (such as dose ranges, mean doses of study drugs) and outcome measures (such as the number of patients who responded to treatment and the number of patients who failed to complete the study by any cause). A double-entry procedure will be employed by two reviewers.

### ***Length of follow up***

In many systematic reviews the ability to provide valid estimates of treatment effect, applicable to the real world, is limited because trials with different durations of follow-up have been combined (Edwards & Anderson, 1999; Geddes et al., 2000; Zimmermann et al., 2002). Clinically, the assessment of efficacy after 6 weeks of treatment or after 16 to 24 weeks or more may lead to wide differences in terms of treatment outcome.

Clinicians need to know whether (and to what extent) treatments work within a clinically reasonable period of time.

One recent systematic review of AD clinical trial data, which investigated the issue of early response to ADs, employed a common definition of early response across all included studies (Taylor et al., 2006). Apart from this review however, no systematic reviews have studied the comparative efficacy of ADs in individuals with major depression employing a common definition of acute response that

includes a pre-defined follow-up duration. In the present review, acute treatment will be defined as an 8-week treatment in both the efficacy and acceptability analyses (Bauer et al., 2002).

If 8-week data are not available, we will use data ranging between 6 to 12 weeks, the time point given in the original study as the study endpoint is given preference.

### ***Quality Assessment***

To assess the quality (internal validity) of trials, we will use predefined criteria based on those developed by the Cochrane Collaboration.

Inadequate concealment undermines the principle of randomization, because participants may then be allocated to a treatment according to prognostic variables rather than by pure chance. Therefore, two independent review authors will independently assess trial quality in accordance with the Cochrane Handbook (Higgins & Green, 2005). This pays particular attention to the adequacy of the random allocation concealment and double blinding (6.11 of the Handbook). Studies will be given a quality rating of A (adequate), B (unclear), and C (inadequate) according to these two items. Studies which will score A or B on these criteria constitute the final list of included studies. In addition, a general appraisal of study quality will be made by assessing key methodological issues such as completeness of follow-up and reporting of study withdrawals.

Where inadequate details of allocation concealment and other characteristics of trials are provided, the trial authors will be contacted in order to obtain further information. If the raters disagree, the final rating will be made by consensus with the involvement (if necessary) of another member of the review group.

Non-congruence in quality assessment will be reported as percentage disagreement.

### ***Comparability of dosages***

In addition to internal and external validity, we will assess the comparability of dosages. Because we could not find any clear definitions about equivalence of dosages among second-generation antidepressants in the published literature, we will use the same roster of low, medium, and high dosages for each drug as Gartlehner and colleagues used in their AHRQ report (Gartlehner et al., 2007) (Table I).

As the authors stated in their report, this classification is based on the interquartile dosing range and doesn't indicate dosing equivalence. This roster will be employed to detect inequalities in dosing that could affect comparative effectiveness.

**Table. Dosing classification based on lower and upper dosing range quartiles**

<b>Drug</b>	<b>Range</b>	<b>Low</b>	<b>Medium</b>	<b>High</b>
Bupropion	150-450 mg/d	< 337.5	337.5-412.5	> 412.5
Citalopram	20-60 mg/d	< 30	30-50	> 50
Duloxetine	60-100 mg/d	< 70	70-90	> 90
Escitalopram	10-30 mg/d	< 15	15-25	> 25
Fluoxetine	20-60 mg/d	< 30	30-50	> 50
Fluvoxamine	50-300 mg/d	< 75	75-125	> 125
Milnacipran	50-300 mg/d	< 75	75-125	> 125
Mirtazapine	15-45 mg/d	< 22.5	22.5-37.5	> 37.5
Paroxetine	20-60 mg/d	< 30	30-50	> 50
Reboxetine	4-12 mg/d	< 5	5-9	> 9
Sertraline	50-200 mg/d	< 75	75-125	> 125
Venlafaxine	75-250 mg/d	< 156.25	156.25-218.75	> 218.75

## STATISTICAL ANALYSIS

Considering that clinical trials of antidepressant drugs are usually small and that data distribution is difficult to assess for studies with small samples, in this review priority will be given to the use and analysis of dichotomous variables both for efficacy and acceptability. When dichotomous efficacy outcomes are not reported but baseline mean and endpoint mean and standard deviation of the depression rating scales (such as HDRS or MADRS) are provided, we will calculate the number of responding patients at 8 weeks (range 6 to 12 weeks) employing a validated imputation method (Furukawa et al., 2005). Even though the change scores give more precision (i.e. narrower 95% CI), we will use for imputation the endpoint scores for the following reasons: (i) standardised mean difference should focus on standard deviation of endpoint scores (standard deviation of change does not represent population variation); (ii) reporting change may represent outcome reporting bias; (iii) we would need to make up more data to impute standard deviation of change scores; (iv) observed standard deviation of change is about the same as observed standard deviation of endpoint. Where outcome data or standard deviations are not recorded, authors will be asked to supply the data. When only the standard error or t-statistics or p values are reported, standard deviations will be calculated according to Altman (Altman, 1996). In the absence of data from the authors, the mean

value of known standard deviations will be calculated from the group of included studies according to Furukawa and colleagues (Furukawa et al., 2006). We will check that the original standard deviations are properly distributed, so that the imputed standard deviation represent the average.

Responders to treatment will be calculated on an intention-to-treat (ITT) basis: drop-outs will always be included in this analysis. When data on drop-outs are carried forward and included in the efficacy evaluation (Last Observation Carried Forward, LOCF), they will be analysed according to the primary studies; when dropouts are excluded from any assessment in the primary studies, they will be considered as drug failures.

### *Synthesis of results*

We will generate descriptive statistics for trial and study population characteristics across all eligible trials, describing the types of comparisons and some important variables, either clinical or methodological (such as year of publication, age, severity of illness, sponsorship, clinical setting).

For each pair-wise comparison between antidepressants, the odds ratio will be calculated with a 95% CI. We will first perform pair-wise meta-analyses by synthesizing studies that compare the same interventions using a random effects model (DerSimonian & Laird, 1986) to incorporate the assumption that the different studies are estimating different, yet related, treatment effects (Higgins & Green, 2006). Visual inspection of the forest plots will be used to investigate the possibility of statistical heterogeneity. This will be supplemented using, primarily, the I-squared statistic. This provides an estimate of the percentage of variability due to heterogeneity rather than a sampling error (Higgins et al., 2003). 95% confidence intervals will be calculated for I-squared, and a P value from a standard test for heterogeneity will be used to assess evidence of its presence.

We will conduct a MTM. MTM is a method of synthesizing information from a network of trials addressing the same question but involving different interventions. For a given comparison, say A versus B, direct evidence is provided by studies that compare these two treatments directly. However, indirect evidence is provided when studies that compare A versus C and B versus C are analyzed jointly. The combination of the direct and indirect into a single effect size can increase precision while randomization is respected. The combination of direct and indirect evidence for any given treatment comparison can be extended when ranking more than three types of treatments according to their effectiveness: every study contributes evidence about a subset of these treatments. We will perform MTM within a Bayesian framework (Ades et al., 2006). This enables us to estimate the probability for each intervention to be the best for each positive outcome, given the results of the MTM. The analysis will be performed using WinBUGS (MRC Biostatistics Unit, Cambridge, U.K., <http://www.mrcbsu.cam.ac.uk/bugs/winbugs/contents.shtml>).

MTM should be used with caution, and the underlying assumptions of the analysis should be investigated carefully. Key among these is that the network is coherent, meaning that direct and indirect evidence on the same comparisons agree. Joint analysis of treatments can be misleading if the network is substantially incoherent, i.e., if there is disagreement between indirect and direct estimates. So, as a first step, we will calculate the difference between indirect and direct estimates in each closed loop formed by the network of trials as a measure of incoherence and we will subsequently examine whether there are any material discrepancies. In case of significant incoherence we will investigate possible sources of it. Incoherence may result as an uneven distribution of effect modifiers across groups of trials that compare different treatments. Therefore, we will investigate the distribution of clinical and methodological variables that we suspect may be potential sources of either heterogeneity or incoherence in each comparison-specific group of trials.

#### *Sensitivity analysis*

Sensitivity analyses will be performed according to the following variables: dose (including only “within the range” studies – see Table) and imputation (including only “without imputation” studies).

## REFERENCES

- Ades AE, Sculpher M, Sutton A, Abrams K, Cooper N, Welton N, Lu G. Bayesian methods for evidence synthesis in cost-effectiveness analysis. *Pharmacoeconomics* 2006;24(1):1-19.
- Altman DG, Bland JM. Detecting skewness from summary information. *BMJ* 1996;313(7066):1200.
- American Psychiatric Association (APA). Diagnostic and Statistical Manual of Mental Disorders (DSM-III). 3rd edition. Washington, DC: American Psychiatric Association, 1980.
- American Psychiatric Association (APA). Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R). 3rd revised edition. Washington, DC: American Psychiatric Association, 1987.
- American Psychiatric Association (APA). Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. Washington, DC: American Psychiatric Association, 1994.
- Bauer M, Whybrow PC, Angst J, Versiani M, Moller HJ; World Federation of Societies Biological Psychiatry Task Force on Treatment Guidelines for Unipolar Depressive Disorders. World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for Biological Treatment of Unipolar Depressive Disorders, Part 1: Acute and continuation treatment of major depressive disorder. *World J Biol Psychiatry* 2002;3(1):5-43.
- Chakrabarti A, Adams CE, Rathbone J, Wright J, Xia J, Wong W, Von Reibnitz P, Koenig C, Baier S, Pfeiffer C, Blatter J, Mantz M, Kloeckner K. Schizophrenia trials in China: a survey. *Acta Psychiatr Scand* 2007;116(1):6-9.
- Cipriani A, Barbui C, Brambilla P, Furukawa TA, Hotopf M, Geddes JR. Are All Antidepressants Really the Same? The Case of Fluoxetine: A Systematic Review. *J Clin Psychiatry* 2006;67(6):850-864.
- Cooper PJ, Murray L. Postnatal depression. *BMJ* 1998;316(7148):1884-6.
- DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177-88

Edwards JG, Anderson I. Systematic review and guide to selection of selective serotonin reuptake inhibitors. *Drugs* 1999;57(4):507-33.

Elliott WJ, Meyer PM. Incident diabetes in clinical trials of antihypertensive drugs: a network meta-analysis. *Lancet* 2007;369(9557):201-7.

Furukawa TA, Cipriani A, Barbui C, Brambilla P, Watanabe N. Imputing response rates from means and standard deviations in meta-analysis. *Int Clin Psychopharm* 2005;20(1):49-52.

Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *J Clin Epidemiol* 2006;59(1):7-10.

Furukawa TA, Watanabe N, Omori IM, et al. Association between unreported outcomes and effect size estimates in Cochrane meta-analyses. *JAMA* 2007;297(5):468-70.

Gartlehner G, Hansen RA, Thieda P, DeVeaugh-Geiss AM, Gaynes BN, Krebs EE, Lux LJ, Morgan LC, Shumate JA, Monroe LG, Lohr KN. Comparative Effectiveness of Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression. Comparative Effectiveness Review No. 7. (Prepared by RTI International-University of North Carolina Evidence-based Practice Center under Contract No. 290-02-0016.) Rockville, MD: Agency for Healthcare Research and Quality. January 2007. Available at: [www.effectivehealthcare.ahrq.gov/reports/final.cfm](http://www.effectivehealthcare.ahrq.gov/reports/final.cfm).

Geddes JR, Freemantle N, Mason J, Eccles MP, Boynton J. Selective serotonin reuptake inhibitors (SSRIs) versus other antidepressants for depression. The Cochrane Library (Cochrane Review). 2000, Issue 2.

Greenberg PE, Kessler RC, Birnbaum HG, Leong SA, Lowe SW, Berglund PA, et al. The economic burden of depression in the United States: how did it change between 1990 and 2000? *J Clin Psychiatry* 2003;64(12):1465-75.

Guy W, Bonato RR. *Manual for the ECDEU Assessment Battery*.2. Chevy Chase, MD: National Institute of Mental Health, 1970.

Hamilton M. (1960) A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;23:56-62.

Hansen AR, Gartlehner G, Lohr KN, Gaynes BN & Carey TS. Efficacy and safety of second-generation antidepressants in the treatment of major depressive disorder. *Ann Int Medicine* 2005;143, 415-26.

Hasselblad V. Meta-analysis of multitreatment studies. *Med Decis Making*. 1998 Jan-Mar;18(1):37-43.

Higgins JP, Whitehead A. Borrowing strength from external trials in a meta-analysis. *Stat Med* 1996;15(24):2733-49.

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-60.

Higgins JP, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions 4.2.5 In: The Cochrane Library, Issue 3, 2005 Chichester, UK: John Wiley & Sons, Ltd., May 2005.

Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions 4.2.6 [updated September 2006]. In: The Cochrane Library, Issue 4, 2006. Chichester, UK: John Wiley & Sons, Ltd.

IMS Health. Press Release: Growth is sustained by new products despite a difficult year [http://www.imshealth.com/ims/portal/front/articleC/0,2777,6599\\_3665\\_44771558,00.html](http://www.imshealth.com/ims/portal/front/articleC/0,2777,6599_3665_44771558,00.html). IMS Reports; 2004.

Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA* 2003;289(23):3095-105.

Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23:3105-3124.

Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med* 2002;21(16):2313-24.

Montgomery SA & Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979;134:382-389.

National Institute for Clinical Excellence (NICE). Depression: management of depression in primary and secondary care. London: National Institute for Clinical Excellence, 2004.

Papakostas GI, Thase ME, Fava M, Nelson JC, Shelton RC. Are Antidepressant Drugs That Combine Serotonergic and Noradrenergic Mechanisms of Action More Effective Than the Selective Serotonin Reuptake Inhibitors in Treating Major Depressive Disorder? A Meta-analysis of Studies of Newer Agents. *Biol Psychiatry* 2007 Jun 21; [Epub ahead of print]

Psaty BM, Lumley T, Furberg CD, Schellenbaum G, Pahor M, Alderman MH, Weiss NS. Health outcomes associated with various antihypertensive therapies used as first-line agents: a network meta-analysis. *JAMA*. 2003 May 21;289(19):2534-44.

Puech A, Montgomery SA, Prost JF, Solles A, Briley M. Milnacipran, a new serotonin and noradrenaline reuptake inhibitor: an overview of its antidepressant activity and clinical tolerability. *Int Clin Psychopharmacol* 1997;12(2):99-108.

Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials, *in press*.

Smith D, Dempster C, Glanville, et al. Efficacy and tolerability of venlafaxine compared with selective serotonin reuptake inhibitors and other antidepressants: a meta-analysis. *Br J Psychiatry* 2002;180:396-404.

Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 2003;326(7387):472.

Taylor MJ, Freemantle N, Geddes JR, Bhagwagar Z. Early onset of selective serotonin reuptake inhibitor antidepressant action: systematic review and meta-analysis. *Arch Gen Psychiatry* 2006;63(11):1217-23.

Watanabe N, Omori IM, Nakagawa A, Cipriani A, Barbui C, McGuire H, Churchill R, Furukawa TA. Mirtazapine versus other antidepressants in the acute-phase treatment of adults with major depression: Systematic review and meta-analysis. *J Clin Psych* (in press).

World Health Organization (WHO). The Ninth Revision of the International Classification of Diseases and Related Health Problems (ICD-9). Geneva: World Health Organization, 1978.

World Health Organization (WHO). The Tenth Revision of the International Classification of Diseases and Related Health Problems (ICD-10). Geneva: World Health Organization, 1992.

Zimmerman M, Posternak MA, Chelminski I. Symptom severity and exclusion from antidepressant efficacy trials. *J Clin Psychopharmacol* 2002;22(6):610-4.