

Agomelatine versus placebo:

A meta-analysis of published and unpublished trials

(Protocol for a systematic review, Ulm, January 17, 2011)

Markus Kösters, Andrea Cipriani, Giuseppe Guaiana, Thomas Becker and Corrado Barbui

Contact Person:

Dr. Markus Kösters
Division Psychiatry II, Ulm University
Bezirkskrankenhaus Guenzburg
Ludwig-Heilmeyer-Str. 2
D-89312 Guenzburg/Germany
Phone: ++49-8221-96-2869

OBJECTIVES

- (1) To determine the acute and long term efficacy of agomelatine in the treatment of unipolar major depression compared to placebo.
- (2) To review the acceptability of agomelatine in comparison to placebo.

METHODS

Types of studies

This systematic review will include only published and unpublished double-blind parallel-group randomised controlled trials. For trials with a crossover design only results from the first randomisation period will be considered.

Types of participants

Studies in adult patients (>18 years) with a primary diagnosis of unipolar major depression according to DSM-III (1), DSM- III-R (2), DSM-IV (3), DSM- IV-TR (4), ICD- 10 (5), Feighner (6) or Research Diagnostic Criteria (7) will be included. Studies including patients with a concurrent primary diagnosis of Axis I or II disorders and antidepressant trials in depressive participants with a serious concomitant medical illness will be excluded.

Types of interventions

Trials comparing agomelatine with placebo as monotherapy in the acute and long term treatment of depression will be included. Only treatment arms within the therapeutic dose range of agomelatine (25-50mg/d) will be included. No restriction in pharmaceutical form or dose regimen (fixed or flexible) will be applied.

Types of outcome measures

Primary outcome

Acute-phase studies: The primary outcome measure for acute phase studies will be the group mean scores at the end of the trial, or group mean change from baseline to endpoint, on Hamilton Depression Rating Scale (HDRS).

Long-term studies: The primary outcome for long term studies will be the proportion of patients who relapsed during the follow-up treatment period. Any definition of depression relapse will be included.

Secondary outcomes

- Group mean scores at the end of the trial, or group mean change from baseline to endpoint, on HDRS, Montgomery-Asberg Depression Scale (MADRS) or Clinical Global Impression Rating scale (CGI), or on any other depression rating scale. When trials reported results from more than one rating scale, we used the HDRS results or, if not available, the MADRS results or, if not available, the results at any other depression rating scale.
- Treatment responders, that is the proportion of patients showing a reduction of at least 50% at the HDRS or MADRS or at any other depression scale (e.g. the Beck Depression Inventory or the CES-D scale; or were 'much or very much improved' (score 1 or 2) at the Clinical Global Impression-Improvement (CGI-I), or proportion of patients who improved using any other pre-specified criterion.
- Treatment remitters, that is the proportion of patients showing remission as defined by: a score of 7 or less at the 17-item HDRS, or 8 or less at longer versions of HDRS; 10 or less at the MADRS; 'not ill or borderline mentally ill' on the CGI-S; or any other equivalent value on a depression scale defined by the authors. Preference will be given to remission rates defined by HDRS or MADRS scores.

Acceptability will be evaluated using the following outcome measures:

- Total number of participants who dropped out during the trial as a proportion of the total number of randomised participants: total dropout rate.
- Number of participants who dropped out due to inefficacy during the trial as a proportion of the total number of randomised participants.
- Number of participants who dropped out due to side effects during the trial as a proportion of the total number of randomised participants.
- Total number of participants experiencing at least some side effects.

Search methods for identification of studies

Literatures searches will be performed in the following databases and article indexes: MEDLINE, CINAHL, EMBASE, PsycInfo, Cochrane Central Register of Controlled Trials (CENTRAL). Controlled vocabulary was utilized where appropriate terms were available, supplemented with keyword searches to ensure accurate and exhaustive results. Language or publication year limits were not applied to any search (Appendix for details).

To supplement the searches of published research, the internet will also be utilized to locate additional clinical trials, unpublished research and/or grey literature. Websites of pharmaceutical companies, clinical trials registers and regulatory agencies will be searched.

Data collection

Selection of studies

Included and excluded studies will be collected following the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA, 8). We will examine all titles and abstracts first, and then obtain full texts of potentially relevant papers. Working independently and in duplicate, two reviewers will read the papers and determined whether they met inclusion criteria. Considerable care will be taken to exclude duplicate publications.

Data extraction and management

Two review authors will use an electronic data extraction form (EPIDATA) to independently extract the data concerning participant characteristics, intervention details and outcome measures. Disagreements will be resolved by discussion and consensus with a third member of the team.

For continuous outcomes, the mean change from baseline to endpoint, the mean scores at endpoint, the SD or standard error (SE) of these values, and the number of patients included in these analyses, will be extracted (9). Data will be extracted preferring the 17-item HDRS over any other version of the HDRS over the MADRS and over the CGI.

For dichotomous outcomes, the number of patients undergoing the randomization procedure, the number of patients rated as responders, remitters, relapsed and the number of patients leaving the study early will be recorded.

Assessment of risk of bias in included studies

The Cochrane risk-of-bias tool will be used (10). This instrument consists of six items. Two of the items assess the strength of the randomization process in preventing selection bias in the assignment of participants to interventions: adequacy of sequence generation and allocation concealment. The third item (blinding) assesses the influence of performance bias on the study results. The fourth item assesses the likelihood of incomplete outcome data, which raise the possibility of bias in effect estimates. The fifth item assesses selective reporting, the tendency to preferentially report statistically significant outcomes. This item requires a comparison of published data with trial protocols, when such are available. The final item refers to other sources of bias that are relevant in certain circumstances, such as, for example, sponsorship bias.

Summary statistics

A double-entry procedure will be employed. Data will be initially entered and analyzed using the Cochrane Collaboration's Review Manager software version 5 (Oxford, England, Cochrane Collaboration), and subsequently entered into a spreadsheet and re-analyzed using the 'metafor' package (11). Outputs were cross-checked for internal consistency.

Continuous data

Despite some critics (12), the HDRS is still the 'gold standard' for assessing antidepressant efficacy in clinical trials. Furthermore, clinical interpretation of results from metaanalysis is greatly simplified if effect sizes are calculated as (raw) mean differences (MD). Consequently, the primary outcome (acute treatment studies) data will be analysed using a mean difference and only scores from the HDRS will be pooled together. As secondary outcome, data will further be analysed using standardised mean differences (SMD), as scores from different depression scales will be pooled. If endpoint data are unavailable, change score data will be used. Where intention-to-treat (ITT) data is available it will be preferred to 'per-protocol analysis'. When only P or standard error (SE) values are reported, standard deviations will be calculated (13).

Dichotomous outcomes

For the primary outcome (long term studies) and for all secondary binary outcomes we will calculate a Mantel-Haenszel risk ratio (RR). Response, remission and relapse on treatment will be calculated using an ITT analysis. Where participants left the study before the intended endpoint, it will be assumed that they have experienced the negative outcome. When outcome data are not reported, trial authors will be asked to supply the data; in case of no response from study authors, we will estimate the number of patients responding to treatment using a validated imputation method (14;15). The robustness of this approach will be checked by sensitivity analysis.

Confidence intervals

A 99% confidence interval (CI) will be calculated for all efficacy estimates according to Barbu and colleagues (16). This approach, instead of a 95% CI approach, will be adopted to have the widest estimate of likely true effect. We set the level of significance at 0.01 as we will make multiple comparisons and we reasoned that only robust differences between treatments should inform clinical practice. In fact, it is more important to avoid the possibility of showing a difference in the absence of a true difference, than to avoid the possibility of not showing a difference in the presence of a true difference. In other words, we give priority to avoid a type I than a type II error (17). Conversely, a 95% CI will be calculated for all tolerability estimates. In terms of tolerability it is more important to avoid the possibility of not showing a difference in the presence of a true difference than to avoid the possibility of showing a difference in the absence of a true difference. In other words, we give priority to avoid a type II than a type I error.

Studies with multiple treatment groups

For dichotomous outcomes, trials comparing different doses of agomelatine with placebo were converted into two-arm trials by summing samples and averaging doses. For continuous outcomes,

means and standard deviations of different dosage arms are combined into a single arm according to the methods described in the Cochrane handbook (10, Chapter 7.7.3.8).

Assessment of heterogeneity

Visual inspection of graphs will be used to investigate the possibility of statistical heterogeneity. This will be supplemented using the I-squared statistic. This provides an estimate of the percentage of variability due to heterogeneity rather than chance alone. Where the I-squared estimate is greater than or equal to 50% we interpreted this as indicating the presence of high levels of heterogeneity (18). Statistical significance of heterogeneity will additionally be tested with chi-square tests, using a threshold of $p < 0.20$ as threshold of statistical significance.

Assessment of publication bias

Funnel plots will be used to investigate publication bias.

Data synthesis and presentation

Continuous and dichotomous outcomes will be analysed using a random-effects-model, as this takes into account any differences between studies even if there is no statistically significant heterogeneity (10). A fixed-effects model will be routinely applied to check for material differences.

A summary of findings (SoF) table will be produced according the methodology described by the GRADE working group (19;20).

Subgroup and sensitivity analysis

The following pre-planned subgroup and sensitivity analyses will be carried out:

- (a) Agomelatine dosing (low dosage: 25 mg/d vs. flexible doses and 50mg/d)
- (b) Publication status (published vs unpublished studies)
- (c) Exclusion of trials with imputed data from responder analyses

Funding

This systematic review is not financially supported by drug companies. We have no conflicts of interest.

References

1. American Psychiatric Association.(1980) Diagnostic and Statistical Manual of Mental Disorders (DSM-III), 3rd Edition. Washington, DC: American Psychiatric Publishing;
2. American Psychiatric Association.(1987) Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R), 3rd Edition revised. Washington, DC: American Psychiatric Publishing;
3. American Psychiatric Association.(1994) Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), 4th Edition. Washington, DC: American Psychiatric Publishing;
4. American Psychiatric Association.(2000) Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR), 4th Edition - revised. Washington, DC: American Psychiatric Publishing;
5. World Health Organization.(1992) The Tenth Revision of the International Classification of Diseases and Related Health Problems (ICD-10). Geneva: World Health Organization,
6. Feighner JP, Robins E, Guze SB, Woodruff RA, Jr., Winokur G, Munoz R. (1972) Diagnostic criteria for use in psychiatric research. Arch Gen Psychiatry 26:57-63.
7. Spitzer RL, Endicott J, Robins E. (1978) Research diagnostic criteria: rationale and reliability. Arch Gen Psychiatry 35:773-82.
8. Moher D, Liberati A, Tetzlaff J, Altman DG. (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 339:b2535.
9. Norman GR. (1989) Issues in the use of change scores in randomized trials. J Clin Epidemiol 42:1097-105.
10. Higgins J, Green S.(2011) Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. <http://www.cochrane-handbook.org/> [last accessed 2012 Dec 14];
11. Viechtbauer W. (2010) Conducting Meta-Analyses in R with the metafor Package. Journal of Statistical Software 36:1-48.
12. Bagby RM, Ryder AG, Schuller DR, Marshall MB. (2004) The Hamilton Depression Rating Scale: has the gold standard become a lead weight? Am J Psychiatry 161:2163-77.
13. Altman DG, Bland JM. (1996) Detecting skewness from summary information. BMJ 313:1200.
14. Furukawa TA, Cipriani A, Barbui C, Brambilla P, Watanabe N. (2005) Imputing response rates from means and standard deviations in meta-analyses. Int Clin Psychopharmacol 20:49-52.
15. Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, et al. (2009) Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. Lancet 373:746-58.
16. Barbui C, Furukawa TA, Cipriani A. (2008) Effectiveness of paroxetine in the treatment of acute major depression in adults: a systematic re-examination of published and unpublished data from randomized trials. CMAJ 178:296-305.

17. Cipriani A, Brambilla P, Furukawa T, Geddes J, Gregis M, Hotopf M, et al. (2005) Fluoxetine versus other types of pharmacotherapy for depression. *Cochrane Database Syst Rev* 4:CD004185.
18. Higgins JP, Thompson SG, Deeks JJ, Altman DG. (2003) Measuring inconsistency in meta-analyses. *BMJ* 327:557-60.
19. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ. (2008) What is "quality of evidence" and why is it important to clinicians? *BMJ* 336:995-8.
20. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y,onso-Coello P, et al. (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 336:924-6.

Appendix

Search strategy

- 1 exp Neurotic Disorders/
- 2 exp Depressive Disorder/
- 3 exp Depression/
- 4 depress\$.ab,hw,ot,sh,ti.
- 5 neurotic disorder\$.ab,hw,ot,sh,ti.
- 6 seasonal affective disorder\$.ab,hw,ot,sh,ti.
- 7 dysthymi\$.ab,hw,ot,sh,ti.
- 8 melanchol\$.ab,hw,ot,sh,ti.
- 9 or/1-8
- 10 randomized controlled trial.pt.
- 11 controlled clinical trial.pt.
- 12 exp Randomized Controlled Trials/
- 13 random allocation.ab,hw,ot,sh,ti.
- 14 exp Random Allocation/
- 15 random\$.ti.
- 16 exp Double-Blind Method/
- 17 exp Single-Blind Method/
- 18 ((singl\$ or doubl\$ or tripl\$ or trebl\$) and (blind\$ or mask\$ or dummy\$)).ab,hw,ot,sh,ti.
- 19 (random\$ and (trial or study)).ab,hw,ot,sh,ti.
- 20 or/10-19
- 21 (agomelatin\$ or valdoxan or thymanax or melitor).mp. [mp=ps, rs, ti, ot, ab, nm, hw, ui, sh, kw, tn, dm, mf, dv, tc, id, tm]
- 22 9 and 20 and 21