

Clinical relevance of findings in trials of CBT for depression

Lepping, Peter; Whittington, R.; Sambhi, R.S.; Lane, S.; Poole, Robert; Leucht, Stefan; Cuijpers, P.; McGabe, R.; Waheed, W.

European Psychiatry

DOI:

[10.1016/j.eurpsy.2017.07.003](https://doi.org/10.1016/j.eurpsy.2017.07.003)

Published: 01/09/2017

Peer reviewed version

[Cyswllt i'r cyhoeddiad / Link to publication](#)

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):

Lepping, P., Whittington, R., Sambhi, R. S., Lane, S., Poole, R., Leucht, S., Cuijpers, P., McGabe, R., & Waheed, W. (2017). Clinical relevance of findings in trials of CBT for depression. *European Psychiatry*, 45, 207-211. <https://doi.org/10.1016/j.eurpsy.2017.07.003>

Hawliau Cyffredinol / General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Title: Clinical relevance of findings in trials of CBT for depression

Authors: Peter Lepping, Richard Whittington, Rajvinder Singh Sambhi, Steven Lane, Rob Poole, Stefan Leucht, Pim Cuijpers, Rhiannah McCabe, and Waquas Waheed

Peter Lepping, MD MRCPsych MSc. ^{1,2,3} (**corresponding author**)

¹ Betsi Cadwaladr University Health Board, Centre for Mental Health and Society, Bangor University, UK

² Mysore Medical College and Research Institute, India

³ Department of Psychiatry, Betsi Cadwaladr University Health Board, North Wales, UK

Wrexham Academic Unit, Centre for Mental Health and Society, Technology Park, Wrexham LL13 7TP, UK, peter.lepping@wales.nhs.uk, Tel: 0044-1978-726752 Fax: 0044-1978-726600

Richard Whittington, BA, PhD, CPsychol, AFBPsS. ⁴

⁴ Department of Health Services Research, University of Liverpool, UK whittington

Rajvinder Singh Sambhi, MD, MRCPsych. ³

Department of Psychiatry, Betsi Cadwaladr University Health Board, North Wales, UK

Steven Lane, PhD MSc BSc (HON). ⁵

⁵ Department of Health Services Research, University of Liverpool, UK

Rob Poole, MB BS, FRCPsych. ¹

Centre for Mental Health and Society, Bangor University, UK

Stefan Leucht, MD. ⁶

⁶ Department of Psychiatry and Psychotherapy, Technische Universität München, Germany

Pim Cuijpers, PhD ⁷

⁷ Department of Clinical Psychology and EMGO Institute for Health and Care Research, VU University and VU University Medical Center Amsterdam, The Netherlands

Rhiannah McCabe, MSc, BSc (hons) ⁸

⁸ Mersey Care NHS Trust, Liverpool UK

Waquas Waheed, MRCPsych, MD. ⁹

⁹ Lancashirecare NHS Trust, Preston and the University of Manchester, UK

Conflict of interest: In the last three years Stefan Leucht has received honoraria for lectures from EliLilly, Lundbeck (Institute), Pfizer, Janssen, BMS, Johnson and Johnson, Otsuka, Roche, SanofiAventis, ICON, Abbvie, AOP Orphan, Servier; for consulting/advisory boards from Roche, Janssen, Lundbeck, EliLilly, Otsuka, TEVA; for the preparation of educational material and publications from Lundbeck Institute and Roche. EliLilly has provided medication for a clinical trial led by SL as principal investigator.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethical Approval: not needed

Acknowledgments:

We thank Rebecca Cannell from the Wrexham Community Mental Health Team for her secretarial support.

ABSTRACT

Cognitive Behavioural Therapy (CBT) is beneficial in depression. Symptom scores can be translated into Clinical Global Impression (CGI) scale scores to indicate clinical relevance. We aimed to assess the clinical relevance of findings of randomised controlled trials (RCTs) of CBT in depression. We identified RCTs of CBT that used the Hamilton Rating Scale for Depression (HAMD). HAMD scores were translated into Clinical Global Impression - Change scale (CGI-I) scores to measure clinical relevance. 170 data sets from 82 studies were included. The mean percentage HAMD change for treatment arms was 53.66%, and 29.81% for control arms, a statistically significant difference. Combined active therapies showed the biggest improvement on CGI-I score, followed by CBT alone. All active treatments, had better than expected HAMD percentage reduction and CGI-I scores. CBT has a clinically relevant effect in depression, with a notional CGI-I score of 2.2, indicating a significant clinical response.

Introduction

1.1 Depressive disorders are common throughout the world. They are by far the most common cause of Years Lost due to Disability in high-, as well as low- and middle-, income countries, in both men and women (Murray & Lopez, 1996). Antidepressants and a range of psychological therapies have been shown to be beneficial in the treatment of depression and are recommended in treatment guidelines (American Psychiatric Association, 2000; National Institute for Clinical Excellence, 2009; Ellis Royal Australian & New Zealand College of Psychiatrists Clinical Practice Guidelines Team for Depression, 2004). People suffering from depression often prefer psychological treatments to medication (Prins et al, 2008). There is variation in individual responses to treatment, including long-term outcome. A range of psychotherapeutic interventions have been developed and refined to improve treatment outcomes (Ellis Royal Australian & New Zealand College of Psychologists Clinical Practice Guidelines Team for Depression, 2004). Clinical trials of psychotherapies suggest that differences between these treatments are small and may have little clinical significance (Cuijpers et al, 2008).

1.2 Previous research has explored the relationship between changes in commonly used symptom rating scales and clinical improvement. Based on data from clinical trials, statistical methods to convert symptom scores into notional Clinical Global Impression scale severity and change (CGI-S and CGI-C/CGI-I) scores have been developed (Leucht et al, 2005; Lepping et al, 2011; Leucht et al, 2013). These two scales quantify clinician's' overall impression of clinical severity and clinical change in participants' psychiatric condition with an aim to bring clinical relevance into trial results. The conversion helps in understanding the relationship between statistically significant results and meaningful clinical improvement.

1.3 Using a published systematic database of randomized trials of psychotherapies for adult depression (Cuijpers et al, 2011), we aimed to assess the clinical relevance of outcomes reported for cognitive behavioural therapy trials in depression. We used a broad definition of CBT therapies, which includes cognitive bibliotherapy, but not behavioural activation therapy. We used a notional CGI translation of scores from the Hamilton Rating Scale for Depression (HAM-D), (Guy W (Ed): ECDEU Assessment Manual for Psychopharmacology.

one of the most commonly used scales in CBT clinical trials (Leucht et al, 2013; Hamilton, 1967).

Material and Methods

2.1 We searched a comprehensive database of randomised controlled clinical trials of psychological therapies for depression (Cuijpers et al 2011; Van't Hof et al, 2011). A detailed description of the development of this database has been published and can be accessed online (www.evidencebasedpsychotherapies.org) (Cuijpers et al, 2008). Systematic review principles were used to include all possible RCT trials using CBT for the treatment of depression.

2.2 The inclusion criteria for our review were:

- Participants: a diagnosis of depression with no psychiatric or physical co-morbidity.
- Interventions: At least one type of psychological therapy classified in the database as CBT¹⁴
- Comparator: any comparator or control. Any arm with a placebo is classed as a control
- Outcome measures: percentage change in mean Hamilton depression rating scale (HAMD/HDRS) score, given directly or calculated from baseline to endpoint data
- Design: randomised controlled trial.
- Reporting:
 - published in a peer reviewed journal and included in the database up to 31.12.2013, the point to which the data base was complete at the time of the study. Books and conference posters were excluded.
 - published in any language.
 - sample size for each study arm reported
 - available as electronic full-text or as paper full-text.

2.3 We obtained the full text version of identified papers. Data were extracted by two reviewers independently (RSS & RW) and differences were resolved by consensus. Leucht and colleagues have demonstrated that changes in HAMD scales can be translated into notional Clinical Global Impressions scale, severity and improvement (CGI-S and CGI-I)

scores (Table 1) (Leucht et al, 2013). The method used for the translation was equi-percentile linking of HAMD-17 and CGI ratings from 43 drug trials in patients with Major Depressive Disorder (MDD) treated with mirtazapine (n=7,131). This method has some limitations, but generates an acceptably robust translation, allowing an objective and reliable estimate of the clinical relevance of published findings. The same or similar methods have been used to evaluate clinical relevance of antipsychotic trial data (Lepping et al, 2011), Transcranial Magnetic Stimulation trial data (Lepping et al, 2014), antidepressant trial data (Moncrieff & Kirsh, 2015), and Cognitive Therapies for schizophrenia trial data (Kontis et al, 2014).

Table 1 about here

2.4 To determine the CGI ratings from the reported HAMD data in our analysis, the mean percentage change from baseline to the last follow-up point was calculated for each study using an Excel spreadsheet. An overall mean percentage change was then calculated for aggregated treatment and control arms, as well as subgroups of particular treatments and control conditions. Hypothesis tests were carried out using the independent sample t-test, at the 5% significance level to assess the statistical significance of mean percentage change between CBT subgroup and the remaining four subgroups. The percentage HAMD change was plotted with CGI-I scores on a graph. The CGI-I score was extracted manually for each point (PL), as the conversion graphs are not linear (Leucht et al, 2013).

2.5 The database classifies a therapy as CBT when cognitive restructuring (the evaluating, challenging, and modifying a patient's dysfunctional beliefs) was one of the core elements of the therapy. The subgroups are classified in accordance with the database. The study arms were categorised in discussion with the research team. The included studies are of diverse quality. They have different degrees of blinding or no blinding at all. Whilst all were RCTs, the recruitment processes were variable and may have relied on volunteers recruited by advertisements. This is in keeping with shortcomings of the psychotherapy literature in general, which are well described (National Institute for Clinical Excellence, 2009). Study arms were classified into five subgroups:

1. CBT alone (n=61 study arms). See above for definition criteria.

2. Other psychological mono-therapy (n=57). This group includes a variety of active psychotherapies, including behavioural activation, provided only one therapy is applied to that group of participants.
3. Pharmacological mono-therapy (n=21). This group includes any antidepressant therapy where only one active drug was used.
4. Combinations of active therapies (n=31). This group includes any combination of therapies, whether psychological or medication.
5. Controls (n=44), this group includes waiting list controls, treatment as usual and placebo.

For a comprehensive list of the included study arms and their categorization see the list of included studies or contact the authors.

Results

3.1 The database was accessed on in December 2013 generating 421 titles (Fig. 1). In the next stage we excluded duplicates and papers with insufficient data. We obtained full-text versions of the remaining 393 papers, of which a further 311 were excluded. The remaining 82 papers were included in the analysis. Overall raw disagreement between the two independent extractors occurred with 13.7% data points. All disagreements were identified and resolved by re-extracting the data. The 82 studies included had between 2 and 10 study arms, yielding 170 datasets relating to a psychological therapy. By adding control arms, a total of 214 datasets were included in the analysis. These datasets included 6330 individual participants. HAMD scores at baseline confirmed that the participants met research criteria for depression, above the accepted threshold score of 8 or higher.

Figure 1 about here

3.2 Table 2 shows the mean percentage change in HAMD scores and the translation into notional CGI-I scores for all five categories. The mean percentage HAMD change for the aggregated treatment arms was 53.66%, and 29.81% for aggregated control arms. This difference was statistically significant ($p < 0.05$). The notional CGI-I score for treatment was 2.25 (1="very much improved", 2="much improved", 3="minimally improved", 4="no change"). For controls it was 3.1.

Table 2 about here

Amongst the different types of psychological and pharmacological treatments, the notional CGI-I score was lowest for the category “combination of active therapies”, indicating the biggest clinical improvement. This was followed by “CBT alone”, followed by “other psychological mono-therapies” and “pharmacological mono-therapies”. Compared to “CBT alone”, the HAMD percentage change of the category “combination of active therapies” showed statistical significance in favour of the “combination of active therapies”. “CBT alone” showed statistically significant superiority to “controls” and “other pharmacological mono-therapies”. It was not superior to “other psychological mono-therapies” (see table 3).

Table 3 about here

3.3 A large sample size was taken as proxy for high quality study. In order to exclude the possibility that the inclusion of many small datasets exaggerated differences in CGI-I scores, we compared the 10 largest datasets with the rest. The 10 largest studies showed a 3% lower HAMD percentage change than the rest of the studies (51.6% versus 54.7%). This equated to a CGI-I score difference of 0.15 (2.3 versus 2.15). We compared findings according to the version of HAMD used. The majority of treatment arms (142) used the 17-item version (HAMD17). Notional CGI-I scores were marginally higher for the 26-item version (used in only one study), but there were no differences between the other three versions. All subgroups had baseline means indicating a diagnosis of depression.

Discussion

4.1 Following the method of previous reviews (Lepping et al, 2011; Lepping et al, 2014; Moncrieff & Kirsch, 2015; Kontis et al, 2014), we translated percentage change in HAMD scores into notional CGI-I scores for psychological therapies in depression. The results indicate that there is a significant and clinically relevant effect of CBT, with a notional CGI-I score of 2.2, indicating that participants were, on average, close to the threshold for being “much improved”. CBT based therapies compared well to other psychological and pharmacological therapies. Only the combination of active therapies showed better results.

The advantage of CBT alone compared to antidepressants alone was statistically significant but below the threshold of clinical relevance (CGI-I difference: 0.15 or 4.36% difference in HAMD score change). CBT was superior to control conditions (CGI-I difference: 0.9, HAMD percentage change difference: 24.21%). The control improvement represents a large placebo effect. The difference between CBT and other psychological mono-therapies was not statistically significant (CGI-I difference: 0.05, HAMD percentage change difference: 1.05%). It therefore appears that the main difference between psychotherapies is how change is achieved, not in the degree of change. Our results suggest important implications for clinicians and service providers, but interpretation of the findings should be cautious. CBT is labour intensive, expensive and requires cooperation from patients over several months. Combinations of active therapies, mainly CBT with antidepressant medication, appears superior to CBT alone, and this may therefore be the most desirable treatment choice.

4.2 The similarity in results for all treatments and combinations, notwithstanding some statistically significant differences, may be a consequence of the small sample sizes in the majority of trials of psychological treatments. Furthermore, there are far more trials of CBT than any other psychological intervention. These shortcomings in the evidence base may militate against capturing differences in clinical effects. It is known that there is significant publication bias in the psychotherapy literature (Cuijpers et al, 2010), and that this may be greater than the corresponding bias in the antidepressant literature. Huhn and colleagues have pointed towards shortcomings in the psychotherapy literature and study designs. They concluded that there are few differences in effect size between medication and psychotherapy. They concluded that whilst effect sizes of psychotherapies versus placebo tended to be higher than those of medication, direct comparisons did not reveal consistent differences. Individual pharmacotherapy trials were more likely to have large sample sizes, blinding, control groups, and intention-to-treat analyses. Psychotherapy trials had lower dropout rates and provided follow-up data (Huhn et al, 2014). They concluded that many pharmacotherapies and psychotherapies are effective, but there is much room for improvement. Because of the multiple differences in the methods used in pharmacotherapy and psychotherapy trials, indirect comparisons of their effect sizes compared with placebo or no treatment are problematic. The same applies to our current analysis. It would be inappropriate to conclude from it that psychotherapy is necessarily more effective than

antidepressants. Well-designed direct comparisons, which are scarce, require public funding, which is scarce (Huhn et al, 2014).

4.3 How do our results compare with other studies? The UK Improving Access to Psychological Therapies project (IAPT) claims a 46.1% recovery rate for those completing treatment. However, few referred patient actually completed the whole course of treatment (<http://www.iapt.nhs.uk/silo/files/iapt-3-year-report.pdf>). Other primary care studies have found that the best remission rates with antidepressants are no higher than 47% (Ramsberg et al, 2012). Even when stringent response criteria are used, such as a CGI-I of 2 (equivalent to a percentage HAMD change of -59%), our results show good response rates in CBT arms for the average participant. In the pharmacological arms, with a mean HAMD change of 49.66%, the average participant responded (using the common industry response criteria of a HAMD reduction of 50%). In these CBT trials, the rate of response to antidepressants far exceeds response rates in drug trials designed to test the effects of antidepressants. Even when we take into account the publication bias in psychotherapy trials, these results are in excess of expectations. It raises the question whether being part of a psychotherapy trial has a large placebo effect. In our analysis the non-specific or placebo effect of being in a psychotherapy trial is 29% reduction in HAMD, which is substantial. Another explanation might be that usual remission criteria have limitations. Zimmerman and colleagues recently found that half of all patients with a HAMD below 7, which is normally considered as remission, did not consider themselves to be in remission and had at least one residual symptom (Zimmerman et al, 2012). The unusual efficacy found in all treatment arms in the psychotherapy literature is a finding that demands further exploration.

4.4 We sourced our studies from a database of psychotherapy trials that finished at the end of 2013. The database we used is exceptional in its meticulous attempt to cover all published trials, but nonetheless we may have missed some studies. Although a large number of datasets were included in the analysis, 52 studies had to be excluded for reporting data in the wrong format. The CBT literature is very heterogeneous. With regards to our analysis, a limitation is that we break randomisation, analyse the different arms individually and then compare them again, instead of first calculating differences between interventions in individual studies and summarising them at the end. This is inherently necessary in our

methodology. Although a limitation, the method nonetheless provides a perspective of the evidence that is useful in drawing conclusions about clinical usefulness.

4.5 Our method of converting HAMD continuous scores into CGI categorical scores is based on translating of psychometrically validated HAMD scores into impressionistic CGI scales by utilising conversion graphs that are not perfectly linear. Based on previous studies, we have reported CGI-I scores to 0.05 differences, to illustrate the proximity of results to thresholds on the CGI-I. Studies used different lengths of treatment, the majority between 3 or 6 months in duration. We have not examined length of treatment as a factor affecting outcome.

4.6 There are variations in the extent to which psychological therapies follow strict CBT procedures. Some modifications of CBT create therapies that are markedly different to the therapy described by Beck. There is no standard measure of fidelity of interventions to the CBT model. There was no alternative but to accept authors' statements that treatments were a form of CBT. Many papers did not describe content of therapy in detail.

Conclusion

5.1 In conclusion, our findings show that there is a significant and clinically relevant effect when using CBT for depression, with a notional CGI-I score of 2.2 indicating a good clinical response. This supports the continuing use of the therapy for this disorder. Our study underlines that statistically significant improvements on specific instruments need to be evaluated in the context of the broader clinical picture.

References

1. American Psychiatric Association (2000) Practice guideline for the treatment of patients with major depressive disorder. 2nd ed. *Am J Psychiatry*. 157(Apr suppl)
2. Cuijpers P, Geraedts AS, van Oppen P, Andersson G, Markowitz JC, van Straten A (2011) Interpersonal psychotherapy for depression: a meta-analysis. *Am J Psychiatry*. Jun;168(6):581-92.
3. Cuijpers P, Smit F, Bohlmeijer E, Hollon SD, Andersson G (2010) Efficacy of cognitive-behavioural therapy and other psychological treatments for adult depression: meta-analytic study of publication bias. *Br J Psych*. Mar 196:173-178
4. Cuijpers P, van Straten A, Andersson G, & van Oppen P (2008) Psychotherapy for depression in adults: A meta-analysis of comparative outcome studies. *Journal of Consulting and Clinical Psychology* 76 (6), 909-922
5. Cuijpers P, van Straten A, Warmerdam L, Andersson G (2008) Psychological treatment of depression: a meta-analytic database of randomized studies. *BMC Psychiatry*. May 16;8:36
6. Ellis P. Royal Australian and New Zealand College of Psychiatrists Clinical Practice Guidelines Team for Depression (2004). Australian and New Zealand clinical practice guidelines for the treatment of depression. *Aust NZ J Psychiatry* 2004; 38:389—407 Guy W (Ed): *ECDEU Assessment Manual for Psychopharmacology*. Washington: US Department of Health, Education, and Welfare; 1976:218–222
7. Guy W (Ed): *ECDEU Assessment Manual for Psychopharmacology*. Washington: US Department of Health, Education, and Welfare; 1976:218–222
8. Hamilton M (1967) Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol*; 6: 278–96
9. Huhn M, Tardy M, Spineli LM, Kissling W, Förstl H, Pitschel-Walz G, Leucht C, Samara M, Dold M, Davis JM, Leucht S (2014) Efficacy of pharmacotherapy and psychotherapy for adult psychiatric disorders: a systematic overview of meta-analyses. *JAMA Psychiatry*. Jun;71(6):706-15. doi: 10.1001/jamapsychiatry.2014.112

10. Kontis D, Theochari E, Tsaltas E (2014) Cognitive therapy for patients with schizophrenia. *Lancet*. Aug; 9941(384):400
11. Lepping P, Sambhi RS, Whittington R, Lane S, Poole R (2011) Clinical relevance of findings in trials of antipsychotics: systematic review. *Br J Psychiatry*. May;198(5):341-5
12. Lepping P, Schönfeldt-Lecuona C, Sambhi RS, Lanka SV, Lane S, Whittington R, Leucht S, Poole R (2014) [A systematic review of the clinical relevance of repetitive transcranial magnetic stimulation](#). *Acta Psychiatr Scand*. Nov;130(5):326-41
13. Leucht S, Fennema H, Engel R, Kaspers M, Lepping P, Szegedi A (2013) Translating HAMD-17 into nominal CGI scores: Clinical implications in patients with Major Depressive Disorder. *J Affect Disord*. Jan 25. doi:pii: S0165-0327(12)00834-8. 10.1016/j.jad.2012.12.001
14. Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel RR (2005) Clinical implications of Brief Psychiatric Rating Scale scores. *Br J Psychiatry*; **187**:366-371
15. Moncrieff J, Kirsch I (2015) Empirically derived criteria cast doubt on the clinical significance of antidepressant-placebo differences. [Contemp Clin Trials](#). Jul;43:60-2
16. Murray CJ, Lopez AD (1996) The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020. *Harvard School of Public Health, Cambridge, MA*. Vol.I (Global Burden of Disease and Injury Series)
17. National Institute for Clinical Excellence (NICE) (2009) Depression: Management of Depression in Primary and Secondary Care (National Clinical Practice Guideline 90). London, NICE
18. Prins MA, Verhaak PF, Bensing JM, van der Meer K (2008) Health beliefs and perceived need for mental health care of anxiety and depression--the patients' perspective explored. *Clin Psychol Rev*. Jul;28(6):1038-58

19. Ramsberg J, Asseburg C, Henriksson M (2012) Effectiveness and cost-effectiveness of antidepressants in primary care: a multiple treatment comparison meta-analysis and cost-effectiveness model. PLoS One; 7(8):e42003.
20. Van't Hof E, Cuijpers P, Waheed W, Stein DJ (2011) Psychological treatments for depression and anxiety disorders in Low- and middle- income countries: a meta-analysis. Afr J Psychiatry (Johannesbg). Jul;14(3):200-7
21. Zimmerman M, Martinez J, Attiullah N, Friedman M, Toba C, Boerescu DA (2012) Symptom differences between depressed outpatients who are in remission according to the Hamilton Depression Rating Scale who do and do not consider themselves to be in remission. J Affect Disord. Dec 15;142(1-3):77-81.
22. <http://www.iapt.nhs.uk/silo/files/iapt-3-year-report.pdf><http://www.iapt.nhs.uk/silo/files/iapt-3-year-report.pdf>

Figure 1: CONSORT
chart

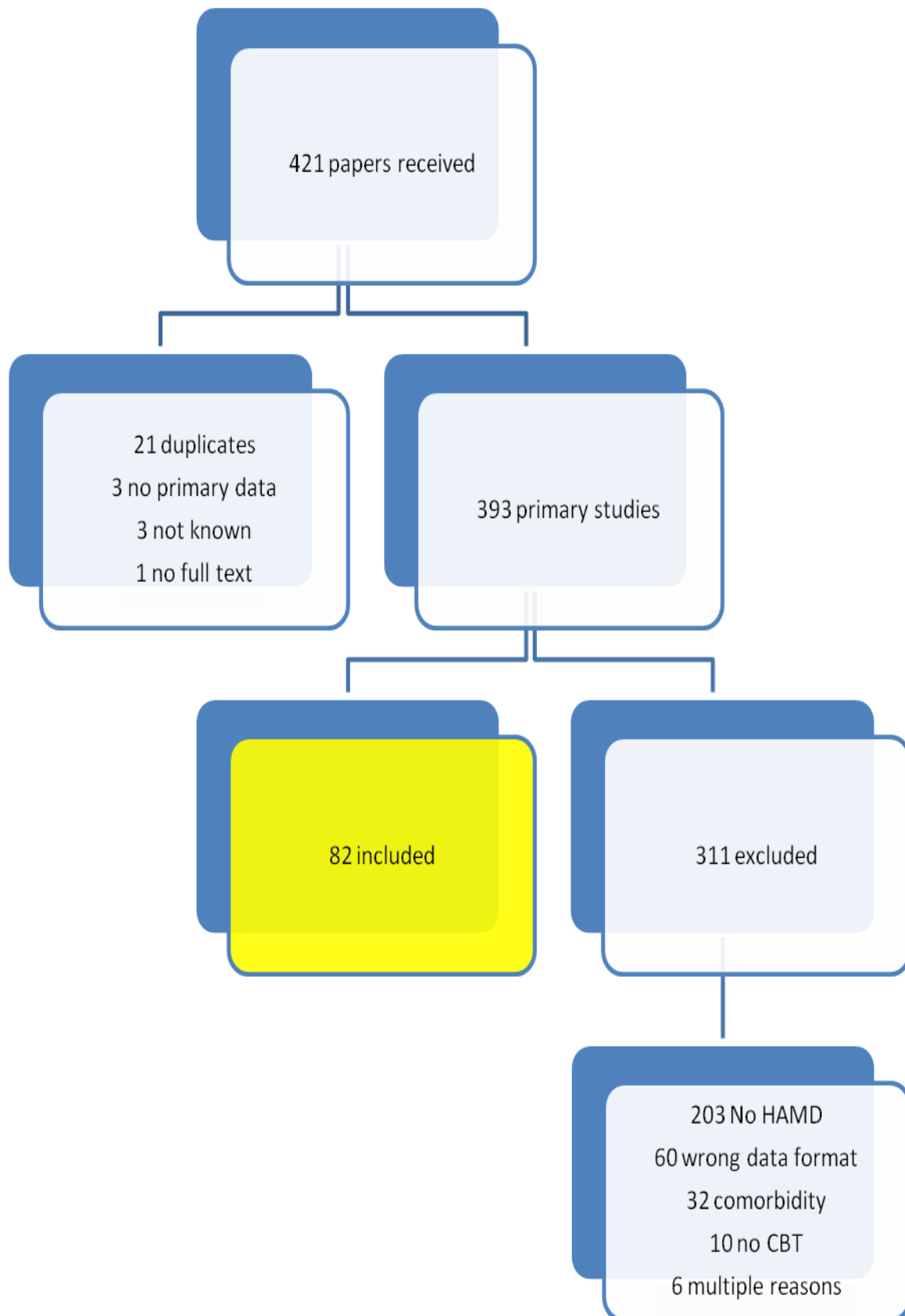


Table 1: CGI translation

HAMD % change	CGI-I	Interpretation
-84%	1	Very much improved
-59%	2	Much improved
-33%	3	Minimally improved
-9%	4	No change
+8%	5	Minimally worse
+27.5%	6	Much worse
+60%	7	very much worse

Table 2: Results for each treatment cluster

Type of treatment	Studies included	Treatment arms included	Total sample size	Combined difference mean (SD)	% change mean (SD)	Notional CGI-I score
Treatment	82	170	4941	10.92 (3.73)	53.66 (15.13)	2.25
CBT alone	49	61	2119	10.13 (2.92)	54.02 (13.68)	2.2
Combination of active therapies	24	31	858	13.45 (4.17)	57.58 (14.25)	2.1
Other psychological monotherapy	43	57	1124	10.59 (4.34)	52.97 (18.16)	2.25
Pharmacological monotherapy	20	21	840	10.88 (3.01)	49.66 (13.92)	2.35
Controls	41	44	1389	5.72 (3.30)	29.81 (17.92)	3.1

Table 3: Hypothesis tests

Comparator 1	Comparator 2	Mean difference in percentage change	T-statistic	Statistically significant difference
CBT	Control	24.21	45.25	Yes
CBT	Combination of active therapies	-3.56	-6.36	Yes
CBT	Other psychological mono-therapy	1.05	1.85	No
CBT	Pharmacological mono-therapy	4.36	7.78	Yes
		T=1.98 implies statistically significant difference between groups		