60 years of placebo-controlled antipsychotic drug trials in acute schizophrenia: Meta-regression of predictors of placebo response

Stefan Leucht, MD, Anna Chaimani, PhD, Claudia Leucht, MD, Maximilian Huhn, MD, Dimitris Mavridis, PhD, Bartosz Helfer, MSc, Myrto Samara, MD, Andrea Cipriani, MD, PhD, John R Geddes, MD, Georgia Salanti, PhD, John M. Davis, MD

Word count body of the text: 2947 words, 2 Tables, 1 Figure, 1 Supplementary File

Corresponding author:

Prof. Stefan Leucht, MD

Department of Psychiatry and Psychotherapy, Technische Universität München

Klinikum rechts der Isar, Ismaningerstr. 22, 81675 Munich, Germany

Tel: +49-89-4140-4249, Fax: +49-89-4140-4888, e-mail: Stefan.Leucht@tum.de

Prof. Andrea Cipriani, MD, PhD, (andrea.cipriani@psych.ox.ac.uk) and Prof. John Geddes, MD, (john.geddes@psych.ox.ac.uk), Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford OX3 7JX, UK

Dr. Claudia Leucht, MD, (Claudia.Leucht@tum.de), Dr. Maximilian Huhn, MD, (Maximilian.Huhn@tum.de), Bartosz Helfer, MSci, (bartosz.helfer@gmail.com), Dr. Myrto Samara (samaramyrto@gmail.com), MD have the same address as Stefan Leucht.

Dr. Anna Chaimani, PhD, (anna.chaimani@parisdescartes.fr): Paris Descartes University; INSERM, UMR1153 Epidemiology and Statistics, Sorbonne Paris Cité Research Center (CRESS), METHODS Team; Cochrane France, Paris, France

Dr. Dimitris Mavridis, PhD, (dimi.mavridis@googlemail.com): Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, University Campus Ioannina 45110, Ioannina, Greece

Prof. Georgia Salanti, PhD (georgia.salanti@ispm.unibe.ch): Institute of Social and Preventive Medicine (ISPM). University of Bern, Finkenhubelweg 11, 3012 Bern, Switzerland

Prof. John M. Davis, MD (davisjm@uic.edu): Psychiatric Institute, University of Illinois at Chicago (mc 912), 1601 W. Taylor St., Chicago, IL 60612, and Maryland Psychiatric Research Center, Baltimore, MD, USA

ABSTRACT

Objective: A recent meta-regression had shown that the degree of placebo response, which has increased over the decades, is the major predictor of drug-placebo differences in antipsychotic drug trials in acutely ill patients with schizophrenia. Drug response, however, had remained stable. In the current meta-regression we explored the factors that are associated with placebo-response.

Method: We searched multiple electronic databases, ClinicalTrials.gov and the FDA website for randomized, placebo-controlled, antipsychotic drug trials in patients with acute exacerbations of schizophrenia. The outcome was the degree of placebo response measured by the BPRS or PANSS change from baseline to endpoint. 26 patient-, design-, and drug-related potential predictors of placebo response were analyzed by univariable and multivariable meta-regressions.

Results: 167 double-blind randomized controlled trials with 28102 participants were included. The mean PANSS change from baseline was 6.25 (95% CI 4.64,7.85). More recent publication year, larger study sample size, more study sites, use of the PANSS rather than the BPRS scale to measure response, shorter wash-out phases, shorter study duration, lower mean age and shorter duration of illness were associated with larger placebo response in univariable analyses. In a multivariable analysis only the number of study participants and mean participant age had an impact on placebo response.

Conclusions: The degree of placebo response is moderated by a number of design and patient-related factors. These explanatory variables of placebo response are only in part identical with those that moderated drug-placebo differences.

Key Words: placebo response, antipsychotics, meta-regression, clinical trials, predictors

1. INTRODUCTION

The efficacy differences between antipsychotic drugs and placebo for acute schizophrenia have decreased over the decades (Leucht et al., 2017), explaining in part why in recent years there have been a number of failed trials where even standard drugs such as haloperidol did not outperform placebo (049, 2010). This decrease in the effect sizes poses a problem for the interpretation of how efficacious antipsychotics really are in practice, and the currently small effect sizes in registrational studies as well as other problems such as high dropout rates (Rabinowitz et al., 2009) pose a major problem for antipsychotic drug development.

A recent meta-regression analysis showed that the degree of placebo response, which has increased over the decades, is the major factor explaining the magnitude of drug-placebo differences (Leucht et al., 2017). Together with industry-sponsorship, it was the only statistically significant predictor of efficacy effect sizes in a multivariable meta-regression. Moreover, when industry-sponsorship, which is probably a composite of various factors (e.g. large sample size, multiple centers and countries, recruitment in professional centers), was removed from a sensitivity analysis, the degree of placebo-response remained the only significant predictor, demonstrating its importance. In contrast, the response in the drug arms of these trials had remained stable over the decades. Therefore, which factors explain *placebo*-response was the major remaining question after our previous analysis of *drug-placebo* differences (Leucht et al., 2017). Previous publications have addressed this issue, as well, but they were all based on approximately 2 times smaller datasets (Agid et al., 2013; Mallinckrodt et al., 2010; Rutherford et al., 2014). When so many more trials are available, the results can easily change, making a new analysis important.

We, therefore, conducted a systematic review and meta-regression analysis of predictors of placebo response. As the aim was to understand which factors explained placebo response in the previous report, we used the same trials compiled in our recently published database of placebo-controlled antipsychotic drug trials in acute schizophrenia, and analysed the same comprehensive list of potential predictors.

2. MATERIAL AND METHODS

We followed the PRISMA guidelines (Liberati et al., 2009) (see checklist in the data supplement S1) and initially published a protocol in PROSPERO which also covered the analysis of placebo response (CRD42013003342, see data supplement S2). The following methods section corresponds to our previous publication with the only major difference that the independent variable was degree of placebo response rather than drug-placebo differences.

2.1 Inclusion/exclusion criteria

2.1.1 Participants

Adults with acute exacerbations of schizophrenia or related disorders (following the Cochrane Schizophrenia Group) were included. We accepted all diagnostic criteria and we also included schizoaffective, schizophreniform, or delusional disorder, because these do not require generally different treatment (Carpenter and Buchanan, 1994). We excluded relapse prevention studies in stable patients receiving maintenance medication (Leucht et al., 2012b), studies in patients with predominant negative symptoms, and studies in patients with major concomitant somatic or psychiatric illness.

2.1.2 Interventions

We included all antipsychotics licensed in at least one country, except clozapine, a more efficacious drug (Leucht et al., 2013a), so that pooling with the other compounds would not have been appropriate *(only one clozapine arm with nine patients had to be excluded on this basis (Honigfeld, 1984) making the impact of this decision negligible)*. We excluded intramuscular formulations, because these are used primarily as sources either for emergency use (short-acting i.m. drugs) or for relapse prevention (long-acting depot drugs). Only the placebo arms of the included studies were used for the current analysis.

Types of studies

Published and unpublished, double-blind, placebo-controlled randomized controlled trials of at least 3 weeks duration (McMahon et al., 2008) were included. Studies with a high risk in sequence generation or

allocation concealment were excluded (Higgins and Green, 2011). We *a priori* excluded Chinese studies due to quality concerns (Woodhead, 2016; Wu et al., 2006). Risk of bias was independently assessed by at least two of the following reviewers (CL, SL, MH, BH) with the Cochrane Collaboration's risk-of-bias tool (Higgins and Green, 2011).

2.2 Search strategy

We searched the Cochrane-Schizophrenia-Group-Controlled-Trials-Register (compiled by regular systematic searches of more than 15 databases, clinical trial registers, the FDA website, hand searches and conference proceedings (Adams et al., 2011), without language restrictions, available to us until version August 2009) with the term "placebo;" and we searched MEDLINE, EMBASE, PsychInfo, Cochrane CENTRAL and ClinicalTrials.gov (last search October 2016, search terms are presented in the online supplement S3), supplemented by screening previous reviews (Adams et al., 2007; Agid et al., 2013; Fenton et al., 2007; Hartung et al., 2005; Joy et al., 2007; Klein and Davis, 1969; Leucht et al., 2009; Leucht et al., 2013a; Matar and Almerie, 2007; Omori and Wang, 2009; Shen et al., 2012).

2.3 Outcomes

The outcome was the mean change from baseline to endpoint of the Positive and Negative Syndrome Scale (PANSS, (Kay and Fiszbein, 1987)) total score. If the PANSS was not available we used the change from baseline to endpoint of the Brief Psychiatric Rating Scale (BPRS, (Overall and Gorham, 1962)) and converted it to the PANSS using a validated method (Leucht et al., 2013b). For this analysis we had applied the equi-percentile linking method to identify corresponding scores of simultaneous BPRS and PANSS ratings in 3767 patients from six antipsychotic drug trials (Leucht et al., 2013b). Higher PANSS change scores means more improvement in the drug group.

2.4 Study selection and data extraction

At least two reviewers among MH, MT, MS and SL independently selected potentially relevant publications from the abstracts found by our search and decided to include studies, and at least two reviewers among CL, MH, BH, MS, MR, SB, MK, PR, TA, NP and SL (see acknowledgement) extracted data in duplicate in Excel sheets. Disagreement was resolved by discussion. Missing data were requested from authors or the sponsoring pharmaceutical companies for all studies published in the last 30 years. We preferably extracted intention-to-treat data and we preferred mixed-effect-model-of-repeated-measurements (MMRM) models over last-observation-carried-forward (LOCF). Missing standard deviations were estimated from test statistics or by using the mean standard deviation of the remaining studies (Furukawa et al., 2006).

2.5 Statistical analysis

We conducted meta-regressions in a frequentist framework with placebo response as the dependent variable. Placebo response was a continuous variable defined as the difference in PANSS/BPRS scale before and after treatment. Predictors of drug-placebo differences could also be predictors of placebo-response we considered as independent variables all study and patient characteristics investigated in our previous article. Only the drug-related explanatory variables (e.g. antipsychotic drug class, mean dose, etc.) were not analysed, because they are not relevant for the placebo groups. The initial choice of predictors has been based on previous evidence (Agid et al., 2013; Furukawa et al., 2015; Mallinckrodt et al., 2010; Rabinowitz et al., 2014; Rutherford et al., 2014) which suggested that these explanatory variables might be relevant. We categorized the explanatory variables into patient-, and study design-related factors, although there were expected overlaps. We first ran univariable meta-regressions exploring separately the effect of each potential explanatory variable. For the multivariable meta-regression model we followed a formal variable selection procedure using the backward stepwise algorithm with removal criterion p=0.15. We monitored how much heterogeneity in placebo response each predictor explains by comparing the heterogeneity of each meta-regression model with the heterogeneity of the model without any covariates.

2.5.1 Patient-related factors

The patient-related factors were: chronicity (Agid et al., 2013) measured by the patients' mean age, duration of illness, duration of the current episode and first episode status (Agid et al., 2013; Rabinowitz et al., 2014); percentage men (Rabinowitz et al., 2014); US American populations versus not/mixed countries (Mattila et al., 2014); severity at baseline (PANSS total score (Furukawa et al., 2014)), in- versus outpatient (Agid et al., 2013); and operationalized criteria (e.g. ICD-10 or DSM-III to IV-R) versus unspecific 'clinical diagnoses.'

2.5.2 Study design-related factors

We analyzed the impact of risk of bias (appropriate versus unclear randomization (Schulz et al., 1995) and allocation concealment methods (Wood et al., 2008), blinding (Wood et al., 2008), and missing outcome data (Higgins and Green, 2011; Porta et al., 2007); study duration (Agid et al., 2013); duration of wash-out (Agid et al., 2013); requirement of a scale-derived minimum of symptoms at baseline (Furukawa et al., 2014); PANSS versus BPRS as a scale; sample size (Egger et al., 1997); number of sites (Agid et al., 2013); percentage of academic sites (Agid et al., 2013); number of medications and arms (Agid et al., 2013); fixed or flexible dosing (Agid et al., 2013); percentage of participants randomized to placebo (Mallinckrodt et al., 2010); and drug company sponsorship of at least one study arm (medication donation alone was not considered company sponsorship (Heres et al., 2006)).

All analyses were performed using Stata 14.2, $p<0.05$ was considered statistically significant.

3. RESULTS

3.1 Description of included studies

The PRISMA (Liberati et al., 2009) flow diagram is presented in the online supplement Figure S1 and a description of the included studies in supplement Table S4. Overall, 167 studies published with 28102 participants met the inclusion criteria, of which 99 studies with 22520 participants (7623 [34%] allocated to the placebo group), which were published from 1969 to 2016 provided data on placebo response. In the studies with data, the patients' mean duration of illness was 14.1 (SD 3.6) years, the mean age 38.7 (SD

4.7) years and the median duration of studies with useable outcomes was 6 weeks (range 3-26 weeks, for the outcome of interest all but one study (26 weeks) lasted ≤12 weeks). There were no studies exclusively examining first-episode patients or treatment-resistant patients. Risk of bias is presented in the online supplement S5. We only included randomised, double-blind trials, but the reports often did not indicate full details about sequence generation, or allocation concealment. Descriptions of methods and success of blinding were frequently insufficient, as well. The data confirmed the high dropout rates in current schizophrenia studies (mean in the placebo groups 52.0%, SD 25.2). Older studies were poorly reported, making it often impossible to extract data (33% of the studies had a high risk of selective reporting). Finally, 64 studies (65%) were sponsored by the manufacturers of one antipsychotic included, 28 (28%) were not primarily industry sponsored and in 7 (7%) studies the sponsor was unclear.

3.2 Explanatory variables of placebo-response – univariable analysis

The mean placebo response in PANSS units was 6.25 (95% CI 4.64,7.85; 99 studies (N) with 7623 participants (n)). This placebo response has increased over the years. The coefficient of 2.74 in Table 1 indicates that a study published 10 years later than another one had, on average, a 2.74 PANSS units higher placebo response. In terms of study design related factors, larger sample size (total number of participants and sites), use of PANSS rather than the BPRS, shorter wash-out phases and shorter study duration were associated with more placebo response. Concerning patient related factors, after excluding an outlier study, the only one restricted to elderly patients with schizophrenia (their mean age was approximately 25 years higher than that of the next oldest population (Tzimos et al., 2008)), lower participant mean age and lower mean duration of illness were associated with higher placebo-response. Moreover, studies conducted exclusively in the US had lower placebo-response than the rest of the studies (Table 1).

3.3 Explanatory variables of placebo response – multivariable analysis

As some significant predictors are related by nature, we made the following choices for the multivariable model: a) We chose publication year to also represent the choice between PANSS versus BPRS to measure response, because the PANSS was only introduced in 1987 so that it could not be used previously b) we chose sample size as representative for the number of sites, because sample size explained more heterogeneity in the univariable analyses. c) we chose mean participant age rather than mean duration ill as a measure of chronicity, because more studies reported this outcome and because age explained more of the heterogeneity in univariable analyses. Then using backward stepwise algorithm for variable selection, total number of participants, age and country were included in the model with the latter not being statistically significant (Table 2a). Further, excluding country from the meta-regression model did not materially change the results (Table 2b). Finally, as it can be argued that increasing the number of sites leads to more variability than increasing the number of patients, we used number of sites rather than participants in a post-hoc sensitivity analysis. The results were virtually identical (Table 2c).

4. DISCUSSION

This meta-regression analysis fills a remaining gap of our previous analysis of predictors of drug-placebo differences in antipsychotic drug trials (Leucht et al., 2017). In the previous analysis, placebo-response was the single strongest predictor of effect sizes. The average placebo response of 6.25 (95% CI 4.64,7.85) was almost one third of the average response in the drug arms 17.45 (95% CI 15.89,19.01) which had remained stable over the decades (Leucht et al., 2017). Thus, it was important to determine the predictors of placebo-response. Our meta-analysis of predictors of placebo response includes two times more studies than all previous ones, and we added several potential explanatory variables that had not been addressed in previous analyses of this kind.

In univariable analyses publication year, number of participants and sites, shorter wash-out phases, shorter study duration, use of the PANSS instead of the BPRS, studies conducted outside the US and less chronicity in terms of lower mean age and duration of illness were associated with more placebo response.

Among these, publication year, the number of sites, study duration and the measures of chronicity were also significant in the next largest analysis of predictors of placebo-response by (Agid et al., 2013). Moreover, in both analyses measures of study quality, the number of arms, the percentage of patients randomised to placebo, and the percentage of men were not significantly associated with placebo response. Given the two times higher number of studies available for our analysis speaks for the robustness of the findings about these explanatory variables. It is also important to note that industry-sponsorship, which was a strong predictor in our analysis of drug-placebo differences (Leucht et al., 2017) and in (Agid et al., 2013), was not significant in the current analysis. Industry sponsorship is probably a composite of various factors (e.g. large sample size, multiple centers and countries, recruitment in professional centers). As there were many companies involved, an analysis of whether individual companies were associated with more placebo-response would have been underpowered. Rutherford et al., 2014 reported a more complex analysis in this context in which not only placebo-controlled trials but also trials that compared antipsychotic drugs with each other were included. Although this difference in approaches makes comparison difficult, there were important similarities e.g. that in their analysis placebo response had also clearly increased over the years and that sample size and trial duration affected this result.

Among the factors that were statistically significant in our results, it is now well established that placebo-response has increased over the years. The average placebo response was 6.25, so that a study conducted in 1970 would have had a worsening in the placebo group of 3 PANSS points, while a study published in 2015 would have an average placebo response of 10 PANSS points. In our previous report we identified multiple design and patient-related factors which have changed over the decades (Leucht et al., 2017). In the current report we identified several ones that were also predictors of placebo response. The finding that longer wash-out phases were associated with less placebo response may reflect that only once patients have been fully washed out from previous treatments they will not respond to placebo. Similarly, a long enough trial duration may be needed for patients to show deterioration rather than improvement under

placebo. Studies using the PANSS had on average higher placebo response than those using the BPRS. It could be that the PANSS is a more sensitive scale than the BPRS, but as the PANSS was introduced only in 1987 (Kay and Fiszbein, 1987), this finding was confounded by publication year. The result in our analysis that studies conducted entirely in the US were associated with less placebo-response than other studies was as surprising as a previous smaller individual-patient data meta-analysis of 21 studies including 5233 patients that found drug-placebo differences were smaller in US American studies (Mattila et al., 2014). As the degree of placebo response was not examined in this previous meta-analysis, we cannot explain the discrepancy, but in our larger meta-analysis of aggregate data, study region was not a significant explanatory variable of drug-placebo differences (Leucht et al., 2017).

However, the only factors that were significant in the multivariable model were participant age and the number of participants. Increasing participant age was associated with less placebo response. We speculate that the symptoms of very chronic patients are quite stable so that they will not change much if patients receive a placebo. The other significant factor, sample size (and the related factor number of sites which was also a significant factor in (Agid et al., 2013)), is a particularly important variable, because it is well-known from other medical fields that small trials lead to larger effect sizes than larger trials (Dechartres et al., 2013; Egger et al., 1997). We chose the number of participants rather than the number of sites in the multivariable model, because it explained more heterogeneity, but arguably more sites may increase variability even more. At the end, we feel that both factors may lead to more variability which would reduce drug-placebo differences (effect size = mean group A – mean group B / standard deviation). In our previous paper about drug-placebo differences we speculated about a vicious circle of increasing sample sizes (and sites), resulting higher variability, smaller effect sizes and which in the next sample size estimation would again mean a higher sample size and more sites (Leucht et al., 2017). The current analysis confirms that extremely high sample sizes should be avoided in placebo-controlled, antipsychotic drug trials.

In the comparison with our previous analysis of antipsychotic drugs versus placebo differences, it is noteworthy that the explanatory variables which were significant in the current analysis of placebo response were not identical with the significant predictors of drug-placebo differences of our previous report (Leucht et al., 2017). Concretely, mean patient age and duration of illness, the duration of the wash-out phase, study duration and country were only significant in the current analysis, while the use of a minimum baseline as an entry criterion, industry sponsorship and the number of medications used, were only significant explanatory variables of drug-placebo differences (Leucht et al., 2017). This finding is important, because it shows that decreasing drug-placebo differences and increasing placebo-response over the years are, at least to some extent, different issues which cannot be explained by exactly the same phenomena.

Our analysis has limitations. In our previous report we reported that there were small trial effects/publication bias in the sense that smaller trials had higher effect sizes than larger ones, and that there may be unpublished smaller trials with smaller effect sizes. Accounting for these effects statistically reduced the effect size from 0.47 to 0.38 (Leucht et al., 2017). The negative association between sample size and placebo response in the current report adds to the interpretation of this finding in that smaller trials had less placebo response. Moreover, if all trials had been identified, different predictors could have resulted. Second, we excluded relapse prevention studies in stable patients receiving maintenance medication, studies in patients with predominant negative symptoms, and studies in patients with major concomitant somatic or psychiatric illness. Other factors may be relevant for placebo response in these populations. (Fraguas et al., 2018) recently reported that more study sites, more study arms and industry sponsorship were significant explanatory variables of placebo response in patients with predominant negative symptoms. Thus, there was partial overlap. Third, we have intentionally not updated our search since the publication of our previous paper (Leucht et al., 2017). We decided not to update the search, because we wanted to understand what predicted placebo response in the studies of our previous analysis of drug-placebo differences. If we had updated the search, it would have been unclear whether differences

in the predictors identified in both reports were due only to the difference in included studies. Nevertheless, our analysis revealed that already four years after the publication of the second-largest analysis of predictors of placebo response (Agid et al., 2013), the number of studies had almost doubled. Fourth, the variable selection for the final multivariable model was based on the backward stepwise algorithm which a) is subject to multicollinearity and b) sometimes leads to over-simplified models by erroneously dropping variables with important contribution to model fit. It is therefore important to not only focus on the multivariable model, but to consider the results of the univariable model, as well. Finally, our multivariable only explained 12% of the heterogeneity. This suggests that there may be further undetected and probably unmeasured predictors of placebo response. For example, monitoring, which has been made mandatory only relatively recently, may be more difficult to conduct if there are many participants. Or there is a speculation about "professional patients/symptomatic volunteers" (Leucht et al., 2012a) who participate in trials, e.g. to benefit from small travel fees. Such participants may report that they have improved to please their doctors, although they received a placebo. But how many such "professional patients" there are is unknown and not recorded. As more trials get published in the future, further analyses will be necessary to determine if the current trend of increasing placebo-response and decreasing effect sizes could stop. Addressing some of the predictors of placebo-response in our report in the methods of future clinical trials could be useful in this regard.

5. REFERENCES

049, S., 2010. A 6-week, double-blind, randomized, fixed dose, parallel-group study of the efficacy and safety of three dose levels of SM-13496 (lurasidone) compared to placebo and haloperidol in patients with schizophrenia who are experiencing an acute exacerbation of symptoms. Center for drug evaluation and research.Application number 200603.Medical review(s).http://www.fga.gov.

Adams, C.E., Awad, G., Rathbone, J., 2007. Chlorpromazine versus placebo for schizophrenia. Cochrane Database of Systematic Reviews(2).

Adams, C.E., Coutinho, E., Davis, J.M., Duggan, L., Essali, A., Fenton, M., Li, C., Jayaram, M., Leucht, S., Tharyan, P., Välimäki, M., 2011. Cochrane Schizophrenia Group, The Cochrane Library. John Wiley & Sons Ltd, Chichester, UK.

Agid, O., Siu, C.O., Potkin, S.G., Kapur, S., Watsky, E., Vanderburg, D., Zipursky, R.B., Remington, G., 2013. Meta-regression analysis of placebo response in antipsychotic trials, 1970-2010. Am J Psychiatry 170(11), 1335-1344.

Carpenter, W.T., Buchanan, R.W., 1994. Schizophrenia. N.Eng.J.Med. 330, 681-690.
Dechartres, A., Trinquart, L., Boutron, I., Ravaud, P., 2013. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. BMJ 346, f2304.
Egger, M., Davey, m.G., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. BMJ 315, 629-634.
Fenton, M., Rathbone, J., Reilly, J., Sultana, A., 2007. Thioridazine for schizophrenia. Cochrane Database Syst Rev(3), CD001944.
Fraguas, D., Diaz-Caneja, C.M., Pina-Camacho, L., Umbricht, D., Arango, C., 2018. Predictors of Placebo Response in Pharmacological Clinical Trials of Negative Symptoms in Schizophrenia: A Meta-regression Analysis. Schizophr Bull.
Furukawa, T.A., Barbui, C., Cipriani, A., Brambilla, P., Watanabe, N., 2006. Imputing missing standard deviations in meta-analyses can provide accurate results. Journal of Clinical Epidemiology 59(1), 7-10.
Furukawa, T.A., Levine, S.Z., Davis, J.M., Samara, M., Leucht, S., 2014. Initial severity and efficacy of antipsychotics: Pooled analysis of six placebo-controlled studies in acute and predominantly negative schizophrenia. JAMA Psychiatry (revision under review).
Furukawa, T.A., Levine, S.Z., Tanaka, S., Goldberg, Y., Samara, M., Davis, J.M., Cipriani, A., Leucht, S., 2015. Initial severity of schizophrenia and efficacy of antipsychotics: participant-level meta-analysis of 6 placebo-controlled studies. JAMA psychiatry 72(1), 14-21.
Hartung, B., Wada, M., Laux, G., Leucht, S., 2005. Perphenazine for schizophrenia. Cochrane Database of Systematic Reviews(1), CD003443.
Heres, S., Davis, J., Maino, K., Jetzinger, E., Kissling, W., Leucht, S., 2006. Why olanzapine beats risperidone, risperidone beats quetiapine, and quetiapine beats olanzapine: An exploratory analysis of head-to-head comparison studies of second-generation antipsychotics. Am.J.Psychiatry 163(2), 185-194.
Higgins, J.P.T., Green, S., 2011. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. Wiley and Sons, Chichester, UK.
Honigfeld, G., 1984. Clozapine: antipsychotic activity in treatment-resistant schizophrenics. Adv.Ther. 1, 77-97.
Joy, C.B., Adams, C.E., Laurie, S., 2007. Haloperidol versus placebo for schizophrenia. Cochrane Database of Systematic Reviews(2).
Kay, S.R., Fiszbein, A., 1987. The positive and negative symptome scale (PANSS) for schizophrenia. Schizophrenia Bulletin 13(2), 261-275.
Klein, D.F., Davis, J.M., 1969. Diagnosis and drug treatment of psychiatric disorders. Williams and Wilkins, Baltimore.
Leucht, S., Arbter, D., Engel, R.R., Kissling, W., Davis, J.M., 2009. How effective are second-generation antipsychotic drugs? A meta-analysis of placebo-controlled trials. Mol.Psychiatry 14(4), 429-447.
Leucht, S., Cipriani, A., Spineli, L., Mavridis, D., Orey, D., Richter, F., Samara, M., Barbui, C., Engel, R.R., Geddes, J.R., Kissling, W., Stapf, M.P., Lassig, B., Salanti, G., Davis, J.M., 2013a. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. Lancet 382(9896), 951-962.
Leucht, S., Hierl, S., Kissling, W., Dold, M., Davis, J.M., 2012a. Putting the efficacy of psychiatric and general medicine medication into perspective: review of meta-analyses. Br J Psychiatry 200(2), 97-106.
Leucht, S., Leucht, C., Huhn, M., Chaimani, A., Mavridis, D., Helfer, B., Samara, M., Rabaioli, M., Bacher, S., Cipriani, A., Geddes, J.R., Salanti, G., Davis, J.M., 2017. Sixty Years of Placebo-Controlled Antipsychotic Drug Trials in Acute Schizophrenia: Systematic Review, Bayesian

Meta-Analysis, and Meta-Regression of Efficacy Predictors. Am J Psychiatry, appiajp201716121358.
Leucht, S., Rothe, P., Davis, J.M., Engel, R.R., 2013b. Equipercentile linking of the BPRS and the PANSS. European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology 23(8), 956-959.
Leucht, S., Tardy, M., Komossa, K., Heres, S., Kissling, W., Salanti, G., Davis, J.M., 2012b. Antipsychotic drugs versus placebo for relapse prevention in schizophrenia: a systematic review and meta-analysis. Lancet 379(9831), 2063-2071.
Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gotzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J., Moher, D., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS.Med. 6(7), e1000100.
Mallinckrodt, C.H., Zhang, L., Prucka, W.R., Millen, B.A., 2010. Signal detection and placebo response in schizophrenia: parallels with depression. Psychopharmacol.Bull. 43, 53-74.
Matar, H.E., Almerie, M.Q., 2007. Oral fluphenazine versus placebo for schizophrenia. Cochrane Database Syst Rev(1), CD006352.
Mattila, T., Wohlfarth, T., Koeter, M., Storosum, J., van den Brink, W., de Haan, L., Leufkens, H., Denys, D., 2014. Geographic variation in efficacy of atypical antipsychotics for the acute treatment of schizophrenia - an individual patient data meta-analysis. European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology 24(7), 1067-1077.
McMahon, R.P., Kelly, D.L., Boggs, D.L., Li, L., Hu, Q., Davis, J.M., Carpenter, W.T., Jr., 2008. Feasibility of reducing the duration of placebo-controlled trials in schizophrenia research. Schizophr Bull 34(2), 292-301.
Omori, I.M., Wang, J., 2009. Sulpiride versus placebo for schizophrenia. Cochrane Database Syst Rev(2), CD007811.
Overall, J.E., Gorham, D.R., 1962. The Brief Psychiatric Rating Scale. Psychol.Rep. 10, 790-812.
Porta, N., Bonet, C., Cobo, E., 2007. Discordance between reported intention-to-treat and per protocol analyses. J Clin Epidemiol 60(7), 663-669.
Rabinowitz, J., Levine, S.Z., Barkai, O., Davidov, O., 2009. Dropout rates in randomized clinical trials of antipsychotics: a meta-analysis comparing first- and second-generation drugs and an examination of the role of trial design features. Schizophr Bull 35(4), 775-788.
Rabinowitz, J., Werbeloff, N., Caers, I., Mandel, F.S., Stauffer, V., Menard, F., Kinon, B.J., Kapur, S., 2014. Determinants of antipsychotic response in schizophrenia: implications for practice and future clinical trials. J Clin Psychiatry 75(4), e308-316.
Rutherford, B.R., Pott, E., Tandler, J.M., Wall, M.M., Roose, S.P., Lieberman, J.A., 2014. Placebo response in antipsychotic clinical trials: a meta-analysis. JAMA psychiatry 71(12), 1409-1421.
Salanti, G., Higgins, J.P., White, I.R., 2006. Bayesian synthesis of epidemiological evidence with different combinations of exposure groups: application to a gene-gene-environment interaction. Stat.Med. 25(24), 4147-4163.
Schulz, K.F., Chalmers, I., Hayes, R.J., Altman, D.G., 1995. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 273, 408-412.
Shen, X., Xia, J., Adams, C.E., 2012. Flupenthixol versus placebo for schizophrenia. Cochrane Database Syst Rev 11, CD009777.
Tzimos, A., Samokhvalov, V., Kramer, M., Ford, L., Gassmann-Mayer, C., Lim, P., Eerdekens, M., 2008. Safety and tolerability of oral paliperidone extended-release tablets in elderly patients

with schizophrenia: a double-blind, placebo-controlled study with six-month open-label extension, American Journal of Geriatric Psychiatry, USA, pp. 31-43.
Wood, L., Egger, M., Gluud, L.L., Schulz, K.F., Juni, P., Altman, D.G., Gluud, C., Martin, R.M., Wood, A.J., Sterne, J.A., 2008. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ 336(7644), 601-605.
Woodhead, M., 2016. 80% of China's clinical trial data are fraudulent, investigation finds. BMJ 355, i5396.
Wu, T.X., Li, Y.P., Liu, G.J., Bian, Z., Li, J., Zhang, J., Xie, L., Ni, J., 2006. Investigation of authenticity of 'claimed' randomized controlled trials (RCTs) and quality assessment of RCT reports published in China. Presented at the XIV Cochrane Colloquium, Dublin, Ireland, October 23-26 2006.

Table 1: Univariable meta-regressions

Explanatory variable	Coefficient	95%CI	p-value	Coefficient corresponds to*	Mean value/reference category of explanatory variable	Weighted mean of placebo response at the mean value/reference category of explanatory variable	95%CI	N, n	Heterogeneity SD of placebo response	% heterogeneity explained
No covariates	-	-	-	-		**6.25**	4.64,7.85	99,7623	8.02	-
Study design related factors										
Publication year[a]	2.74[a]	1.60, 3.88	<0.001	10-year increase	2001	6.24	4.80,7.70	99,7623	7.26	9.5%
Number of total participants[a]	0.16[a]	0.07, 0.24	<0.001	10 patients more	236	6.24	4.74,7.75	99,7623	7.55	5.9%
Number of sites[a]	1.13[a]	0.45,1.81	0.001	10-site increase	30	6.02	4.39,7.66	90,6736	7.80	2.7%
Minimum duration of the wash-out phase[a]	-1.41[a]	-2.33, -0.49	0.003	10-day increase	8	5.06	3.39,6.73	84,6179	7.70	4.0%
Study duration[a]	-8.80	-13.97, -3.63	0.001	10-week increase	6 weeks	6.58	5.11,8.05	94,7558	7.17	10.6%
Scale (PANNS or BPRS)[a]	5.96[a]	2.66,9.26	0.001	BPRS	PANSS	2.19	-0.55,4.94	98,7596	7.57	5.6%
Number of medications	-2.19	-5.64,1.25	0.209	More than one drugs	2 drugs	7.75	4.90,10.61	99,7623	8.00	0.2%
Baseline severity entry minimum score	-1.17	-5.57,3.23	0.597	With Min. entry score	without entry score	7.91	3.85,11.98	75,6156	6.76	15.7%
Industry sponsored drug or not	1.88	-1.49,5.24	0.271	Sponsored	non-	5.11	2.30,7.9	92,75	7.48	6.7%

					sponsored		2	08		
Percentage of academic sites	0.65	-1.35,0.065	0.071	10% increase	56%	4.63	2.15,7.12	53,3019	9.01	0%
Number of arms	2.21	-3.12,7.54	0.412	More than two arms	2 arms	4.26	-0.79,9.31	99,7623	8.04	0%
Percentage randomized to placebo	-0.99	-3.06,1.08	0.345	10% increase	28.0%	6.25	4.64,7.85	99,7623	8.03	0%
Fixed versus flexible dosing	-0.82,	-4.52,2.89	0.663	Flexible dose	Fixed dose	6.45	4.59,8.31	99,7623	8.06	0%
Randomization	-0.13	-3.36,3.09	0.934	Unclear risk	low risk	6.32	3.96,8.68	99,7623	8.06	0%
Allocation concealment	-1.11	-4.57,2.36	0.528	Unclear risk	low risk	7.00	4.14,9.87	99,7623	8.05	0%
Blinding	1.47	-1.75,4.68	0.367	Unclear or high risk	low risk	5.56	3.37,7.75	99,7623	8.03	0%
Risk of bias due to missing outcome data	-2.70	-6.27,0.88	0.138	Unclear or high risk	low risk	6.98	5.12,8.85	99,7623	7.97	0.6%
Patient related factors										
Average age[a,c]	-4.72[a]	-9.31, -0.13	0.044	10-year increase	38	6.27	4.69,7.84	98,7445	7.73	3.6%
Duration ill[a,c]	-9.92[a]	-18.16, -1.68	0.149	10-year increase	14	6.27	4.00,8.54	56,4683	8.47	
Country [a]	3.64[a]	0.48,6.81	0.024	Non-USA or mixed study	USA	4.18	1.80,6.56	99,7623	7.85	2.1%
Baseline severity measured by PANSS[b]	2.52	-0.59, 5.63	0.111	10-unit PANSS increase	94.6	6.99	5.40,8.58	99,7623	7.37	8.1%
Operationalized criteria or not [a]	-6.18	-10.97,-1.39	0.012	No operationalized criteria	Op.criteria	6.94	5.27,8.62	85,7064	7.82	2.5%
Percentage of men	0.20	-0.67,1.07	0.649	10% increase	67.3%	6.50	4.85,8.15	87,7047	7.75	3.4%

[a]Statistically significant explanatory variables, N = number of studies, n= number of patients, PANSS = Positive and Negative Syndrome Scale, BPRS = Brief Psychiatric Rating Scale. [b]When baseline severity was measured by the Clinical Global Impression the result was also not significant (B=4.21 (-2.29, 10.71), B corresponds to 1 CGI unit increase. [c]These analyses were performed after the exclusion of an outlier study (Tzimos et al., 2008). It was the only study which was restricted to elderly patients with schizophrenia. The mean patient age was approximately 25 years higher than that of the next oldest study population.
Coefficient = coefficient of the meta-regression. 95% CI = 95% confidence interval. SD = standard deviation.
"Coefficient corresponds to" = for example publication year: a 10-year increase in publication year on the average increases the placebo response by 2.74 PANSS units.

#Not enough data were available for the variable number of patients with a first episode and there were too few data for ‘duration of the current episode.’ The vast majority of studies included only inpatients. Therefore these parameters could not be analyzed in a meaningful way.
* Coefficient corresponds to a particular increase in the explanatory variable (for continuous characteristics) or to the non-reference category for binary characteristics.

Table 2a: Multivariable meta-regression model resulted from backward stepwise algorithm

Explanatory variable	Coefficient	95% CrI	p-value	Coefficient corresponds to	Interpretation
Total participants[a]	0.15	0.07, 0.24	0.001	10 participants more	For every 10-participants increase in the sample size the average placebo response increases by 0.15 PANSS/BPRS units
Average age[a]	-6.23	-10.46, -2.00	0.004	10 years increase	For every 10 years increase in the average participant's age, the average placebo response decreases by 6.23 PANSS/BPRS units
Country	2.57	-0.52, 5.67	0.102	Non-USA or mixed study (vs USA)	For non-USA or mixed studies in comparison with USA studies, the average placebo response is larger by 2.57 PANSS/BPRS units

Summary of the model: 95 studies with 22300 participants, heterogeneity standard deviation 7.02, the model explained 12% of the heterogeneity

[a]Statistically significant explanatory variables.

Table 2b: Multivariable meta-regression model with age and total participants as predictors

Explanatory variable	Coefficient	95% CrI	p-value	Coefficient corresponds to	Interpretation
Total participants	0.18	0.09, 0.26	<0.001	10 participants more	For every 10-participants increase in the sample size the average placebo response increases by 0.18 PANSS/BPRS units
Average age[a]	-5.94	-10.20, -1.69	0.007	10 years increase	For every 10 years increase in the average participant's

					age, the average placebo response decreases by 5.94 PANSS/BPRS

Summary of the model: 95 studies with 22300 participants, heterogeneity standard deviation 7.09, the model explained 12% of the heterogeneity
[a]Statistically significant explanatory variables.

Table 2c: Multivariable meta-regression model with age and number of sites as predictors

Explanatory variable	Coefficient	95% CrI	p-value	Coefficient corresponds to	Interpretation
Number of sites	1.38	0.71, 2.04	<0.001	10 sites more	For every 10- sites increase the average placebo response increases by 1.38 PANSS/BPRS units
Average age[a]	-6.63	-11.13, -2.14	0.004	10 years increase	For every 10 years increase in the average participant's age, the average placebo response decreases by 6.63 PANSS/BPRS

Summary of the model: 86 studies with 20500 participants, heterogeneity standard deviation 7.27, the model explained 9% of the heterogeneity
[a]Statistically significant explanatory variables.

Legends of Figures

Figure panel 1: Explanatory variables of placebo-response – univariable meta-regressions

The figures in this panel correspond to the following explanatory variables: a) Publication year, b) Number of participants, c) Number of sites, d) Mean age in years, e) Mean duration of illness in years, f) Study duration in weeks, g) Minimum duration of the wash-out phase in days, h) Scale used – PANSS or BPRS, i) Study region, j) Mean PANSS total score at baseline, k) Percentage of men, l) Percentage of academic sites, m) Percentage of patients randomised to placebo, n) minimum scale-derived severity threshold as inclusion criterion, o) number of drugs, p) number of arms, q) Operationalised critera or not, r) industry-sponsored study or not, s) risk of bias concerning randomization method, t) risk of bias

concerning allocation concealment, u) risk of bias concerning blinding, v) risk of bias concerning missing outcomes

The numbers in square brackets describe to what the coefficient refers. For example, in Figure 1a publication year: "B=2.74 (1.60, 3.88) [10 years increase]" means a study that was conducted 10 years later had on average 2.74 (95% confidence interval 1.60 to 3.88) PANSS points higher placebo response. Or, as an example for a dichotomous explanatory variable: B = 5.96 (2.66, 9.26) [PANSS instead of BPRS] means that a study using the PANSS had on the average a 5.96 (95% confidence interval 2.66 to 9.26) PANSS points higher placebo response.

The explanatory variables are statistically significant if the 95% confidence interval does not include 1

[1] results without one outlier which was the only study restricted to elderly people with schizophrenia who had a mean age of 70 years, 20 years more than the next oldest study population,[2] this meta-regression was also statistically significant when the only outlier study of a duration of 26 weeks was excluded

Role of the funding source

The funding sources had no role in study design, data collection, data analysis, data interpretation, writing of the report, or the decision to submit the paper for publication.

Contributors

SL, MH, JMD and GS were involved in designing the review. MH, MT and SL identified and acquired reports of trials. CL, MH, BH, MS, MR, SB, MK, PR, TA, NP and SL extracted data. SL and CL contacted authors of trials and pharmaceutical industries for additional information. SL, ACh, DM, GS and JMD analysed and interpreted the data. ACi, and JRG contributed to the interpretation of the data. SL and ACh drafted the manuscript and all other authors critically reviewed the manuscript for important intellectual content.

Conflict of interest

Stefan Leucht has received honoraria for consulting from LB Pharma, Lundbeck, Otsuka, TEVA, LTS Lohmann, Geodon Richter, Recordati, Boehringer Ingelheim, and for lectures from Janssen, Lilly, Lundbeck, Otsuka, SanofiAventis and Servier. Claudia Leucht is Stefan Leucht's spouse. Maximilian Huhn received lecture honoraria from Janssen and Lundbeck. Andrea Cipriani was expert witness for Accord Healthcare for a patent issue about quetiapine extended release. The other authors have no conflicts of interest to declare. Andrea Cipriani is supported by the National Institute for Health Research (NIHR) Oxford Cognitive Health Clinical Research Facility. John R Geddes is an NIHR Senior Investigator. All other authors declare no competing interests.

Acknowledgement

The meta-analysis was supported by the German Federal Ministry for Education and Research (Bundesministerium für Bildung und Forschung, BMBF) Grant: FKZ 01KG1115. Andrea Cipriani is supported by the NIHR Oxford Cognitive Health Clinical Research Facility. We thank Samantha Roberts for her help in the literature search, Magdolna Tardy, MSc, for her help in study selection and Marc Krause, Philipp Rothe, MD, Thomas Arndt and Matteo Rabaioli, Susanne Bächer, Natalie Peter for their help in data extraction. They had all been trained specifically for this purpose. Ole Andreassen, Guy Chouinard, Michael Jann, Herbert Meltzer, Ofer Agid, Bret Rutherford, Merck, Dainippon Sumitomo, Novartis, Sunovion, EliLilly, Johnson&Johnson and Pfizer for replying to requests about their studies. We also thank the many authors who sent data for our previous publications on antipsychotic drugs which could be used again here.

Figure(s)

Figure 1: Moderators of placebo response – univariable meta-regressions

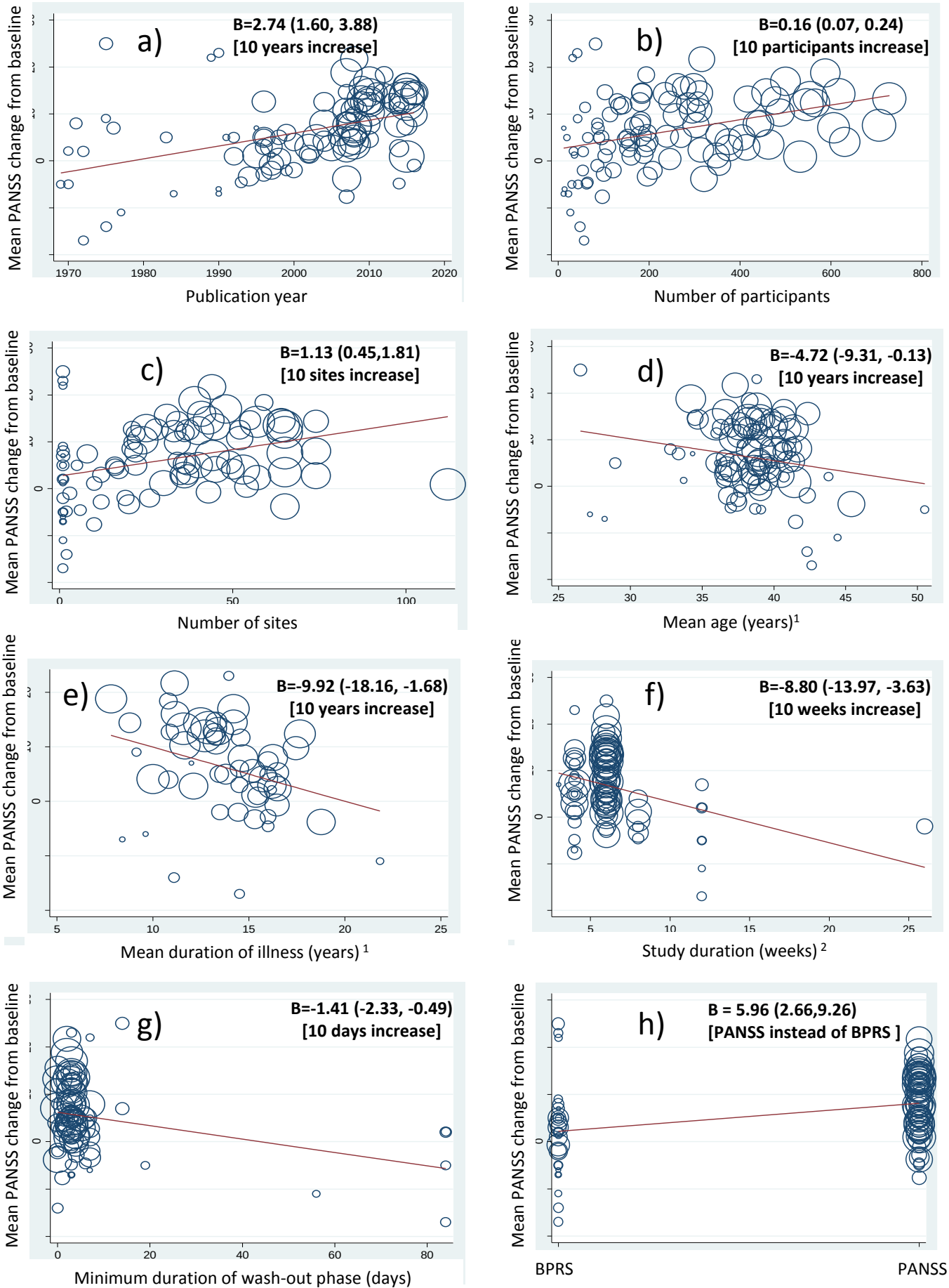
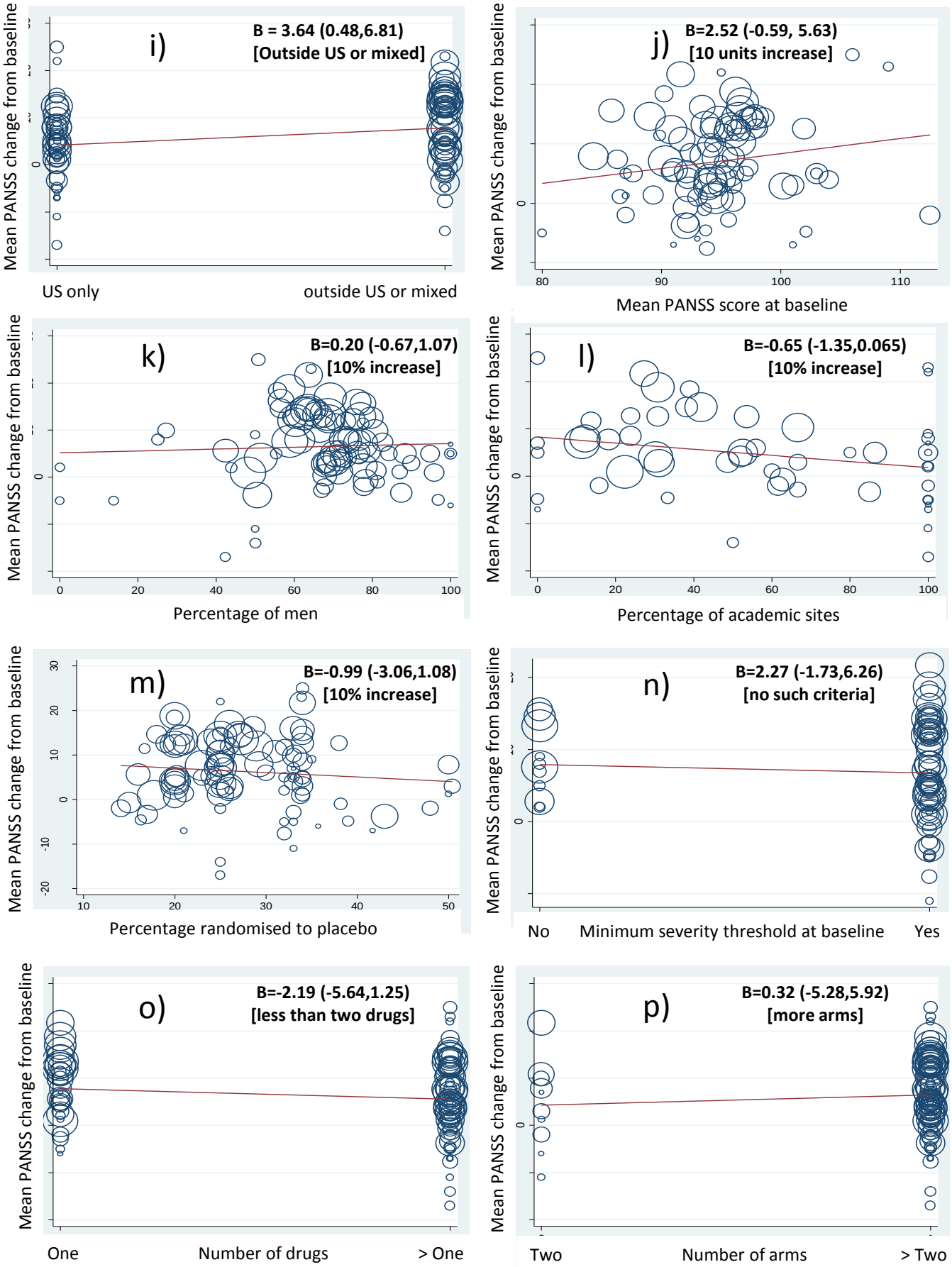
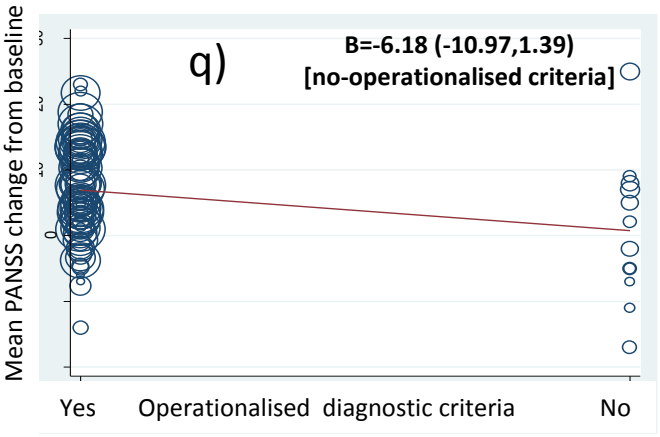
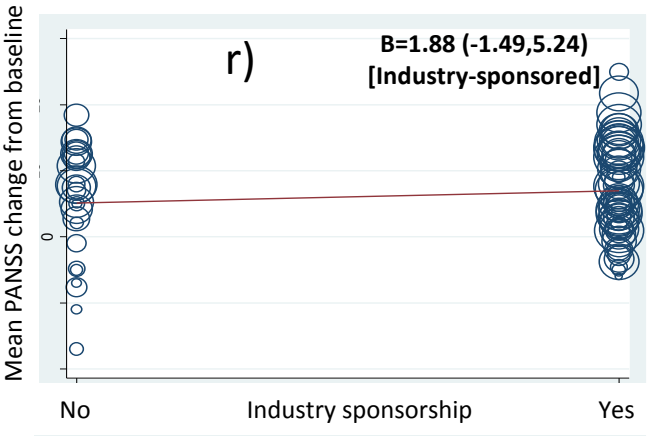
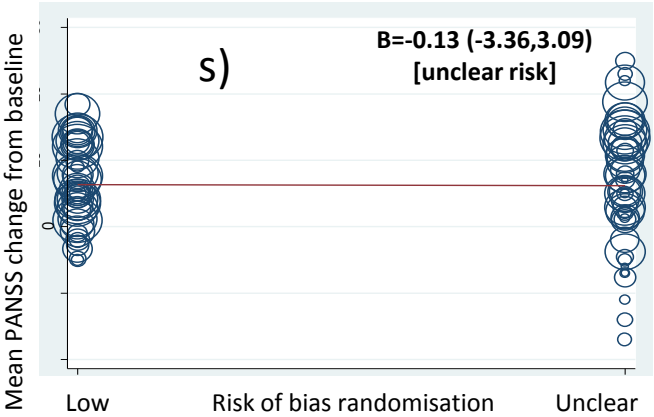


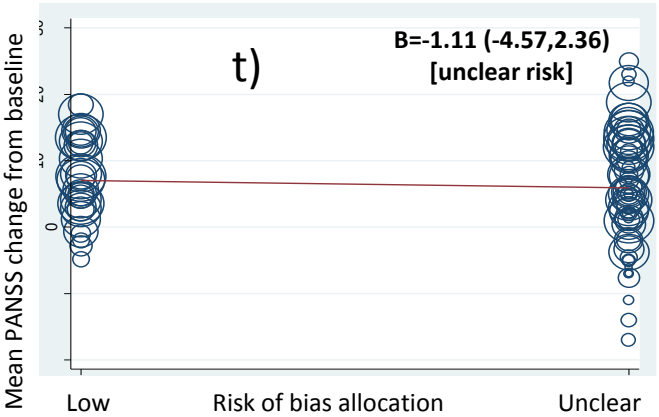
Figure 1: Moderators of placebo response – univariable meta-regressions (continued)

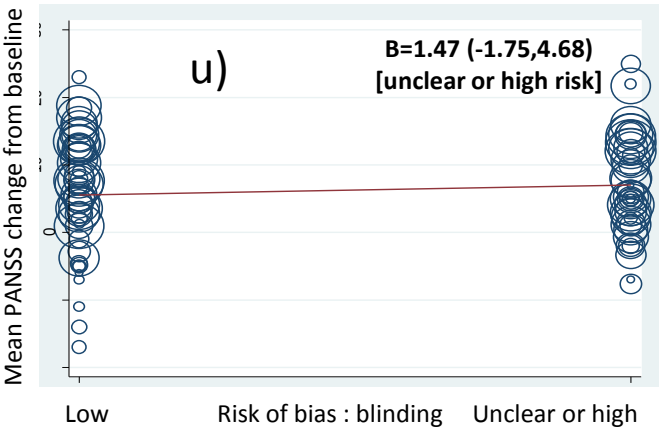


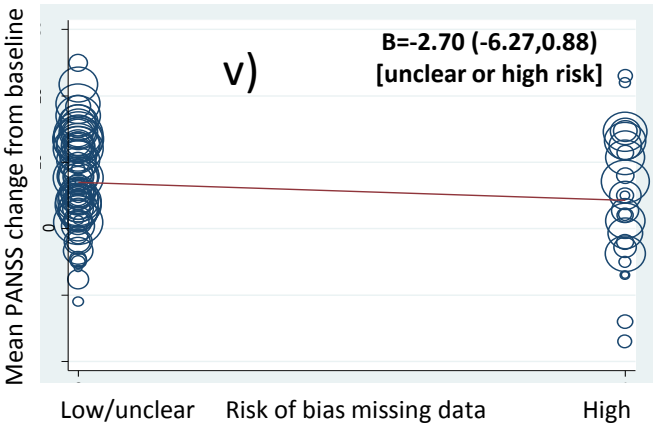

q)
B=-6.18 (-10.97,1.39)
[no-operationalised criteria]
Mean PANSS change from baseline
Yes
Operationalised diagnostic criteria
No

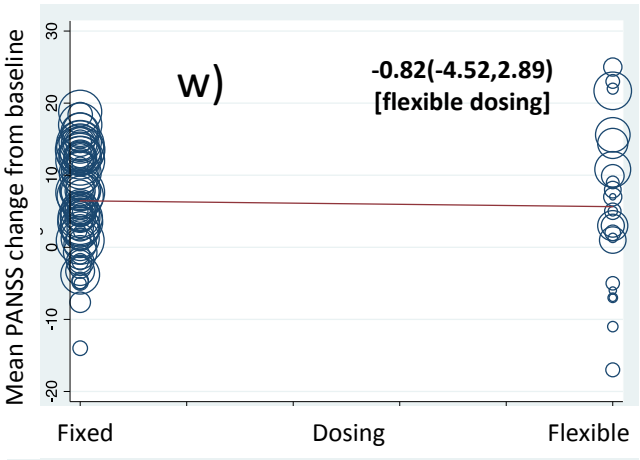

r)
B=1.88 (-1.49,5.24)
[Industry-sponsored]
Mean PANSS change from baseline
No
Industry sponsorship
Yes


s)
B=-0.13 (-3.36,3.09)
[unclear risk]
Mean PANSS change from baseline
Low
Risk of bias randomisation
Unclear


t)
B=-1.11 (-4.57,2.36)
[unclear risk]
Mean PANSS change from baseline
Low
Risk of bias allocation
Unclear


u)
B=1.47 (-1.75,4.68)
[unclear or high risk]
Mean PANSS change from baseline
Low
Risk of bias : blinding
Unclear or high


v)
B=-2.70 (-6.27,0.88)
[unclear or high risk]
Mean PANSS change from baseline
Low/unclear
Risk of bias missing data
High


w)
-0.82(-4.52,2.89)
[flexible dosing]
Mean PANSS change from baseline
Fixed
Dosing
Flexible